

CAMBRIDGE
CRITICAL
GUIDES

KANT'S

Lectures on Ethics

A Critical Guide

Edited by
Lara Denis and Oliver Sensen

KANT'S *LECTURES ON ETHICS*

This is the first book devoted to an examination of Kant's lectures on ethics, which provide a unique and revealing perspective on the development of his views. In fifteen newly commissioned essays, leading Kant scholars discuss four sets of student notes reflecting different periods of Kant's career: Herder (circa 1762–1764), Collins (roughly 1774–1777), Mrongovius (1784–1785), and Vigilantius (1793–1794). The essays cover a diverse range of topics, from the relation between Kant's lectures and the Baumgarten textbooks, to obligation, virtue, love, the highest good, freedom, the categorical imperative, moral motivation, and religion. Together they will provide the reader with a deeper and fuller understanding of the evolution of Kant's moral thought. The volume will be of interest to a range of readers in Kant studies, ethics, political philosophy, religious studies, and history of ideas.

LARA DENIS is Professor of Philosophy at Agnes Scott College. She is the editor of *Kant's Metaphysics of Morals: A Critical Guide* (Cambridge, 2010), and has published numerous essays in journals and books, including in *Kant-Studien*, *Kantian Review*, and *Philosophy and Phenomenological Research*, and *Kant's Tugendlehre: A Comprehensive Commentary* (2013), *Perfecting Virtue: New Essays in Kantian Ethics and Virtue Ethics* (Cambridge, 2011), and *The Cambridge Companion to Kant and Modern Philosophy* (Cambridge, 2006).

OLIVER SENSEN is Associate Professor and Director of Graduate Studies in Philosophy at Tulane University. He is the author of *Kant on Human Dignity* (2011), editor of *Kant on Moral Autonomy* (Cambridge, 2012), and co-editor of *Kant's Tugendlehre* (2013).

CAMBRIDGE CRITICAL GUIDES

Titles published in this series:

Hegel's *Phenomenology of Spirit*

EDITED BY DEAN MOYAR AND MICHAEL QUANTE

Mill's *On Liberty*

EDITED BY C. L. TEN

Kant's *Idea for a Universal History with a Cosmopolitan Aim*

EDITED BY AMÉLIE OKSENBURG RORTY AND JAMES SCHMIDT

Kant's *Groundwork of the Metaphysics of Morals*

EDITED BY JENS TIMMERMANN

Kant's *Critique of Practical Reason*

EDITED BY ANDREWS REATH AND JENS TIMMERMANN

Wittgenstein's *Philosophical Investigations*

EDITED BY ARIF AHMED

Kierkegaard's *Concluding Unscientific Postscript*

EDITED BY RICK ANTHONY FURTAKE

Plato's *Republic*

EDITED BY MARK L. MCPHERRAN

Plato's *Laws*

EDITED BY CHRISTOPHER BOBONICH

Spinoza's *Theological-Political Treatise*

EDITED BY YITZHAK Y. MELAMED AND MICHAEL A. ROSENTHAL

Aristotle's *Nicomachean Ethics*

EDITED BY JON MILLER

Kant's *Metaphysics of Morals*

EDITED BY LARA DENIS

Nietzsche's *On the Genealogy of Morality*

EDITED BY SIMON MAY

Kant's *Observations and Remarks*

EDITED BY RICHARD VELKLEY AND SUSAN SHELL

Augustine's *City of God*

EDITED BY JAMES WETZEL

Descartes' *Meditations*

EDITED BY KAREN DETLEFSEN

Kant's *Religion within the Boundaries of Mere Reason*

EDITED BY GORDON MICHALSON

Kant's *Lectures on Anthropology*

EDITED BY ALIX COHEN

Kierkegaard's *Fear and Trembling*

EDITED BY DANIEL CONWAY

Kant's *Lectures on Ethics*

EDITED BY LARA DENIS AND OLIVER SENSEN

KANT'S
Lectures on Ethics

A Critical Guide

EDITED BY

LARA DENIS

Agnes Scott College

and

OLIVER SENSEN

Tulane University



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107036314

© Cambridge University Press 2015

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2015

Printed in the United Kingdom by Clays, St Ives plc

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data

Kant's lectures on ethics : a critical guide / edited By Lara Denis, Agnes Scott College, and Oliver Sensen, Tulane University.

pages cm. – (Cambridge critical guides)

Includes bibliographical references and index.

ISBN 978-1-107-03631-4

1. Kant, Immanuel, 1724–1804. 2. Ethics. I. Denis, Lara, 1969– editor.

B2799.E8K38 2014

170–dc23

2014023805

ISBN 978-1-107-03631-4 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

<i>Contributors</i>	page ix
<i>Foreword</i>	xiii
<i>J. B. Schneewind</i>	
<i>Translations and abbreviations</i>	xvi
 Introduction	 I
<i>Lara Denis and Oliver Sensen</i>	
 PART I THE SOURCES	 13
1 Kant's lectures on ethics and Baumgarten's moral philosophy	15
<i>Stefano Bacin</i>	
2 Herder: religion and moral motivation	34
<i>Patrick R. Frierson</i>	
3 Collins: Kant's proto-critical position	51
<i>Manfred Kuehn</i>	
4 Mrongovius II: a supplement to the <i>Groundwork of the</i> <i>Metaphysics of Morals</i>	68
<i>Jens Timmermann</i>	
5 Vigilantius: morality for humans	84
<i>Robert B. Louden</i>	
 PART II PRACTICAL PHILOSOPHY	 101
6 Ancient insights in Kant's conception of the highest good	103
<i>Stephen Engstrom</i>	

7	Kant's history of ethics <i>Allen W. Wood</i>	120
8	Moral obligation and free will <i>Oliver Sensen</i>	138
9	The elusive story of Kant's permissive laws <i>B. Sharon Byrd</i>	156
10	On the logic of imputation in the Vigilantius lecture notes <i>Joachim Hruschka</i>	170
PART III ETHICS		185
11	Freedom, ends, and the derivation of duties in the Vigilantius notes <i>Paul Guyer</i>	187
12	Proper self-esteem and duties to oneself <i>Lara Denis</i>	205
13	Virtue, self-mastery, and the autocracy of practical reason <i>Anne Margaret Baxley</i>	223
14	Love <i>Jeanine Grenberg</i>	239
15	Love of honor, emulation, and the psychology of the devilish vices <i>Houston Smit and Mark Timmons</i>	256
<i>Works cited</i>		277
<i>Index</i>		288

Contributors

STEFANO BACIN is currently Lecturer in Moral Philosophy at the Università San Raffaele (Milan), and has been an Alexander von Humboldt Fellow at the Johannes Gutenberg-Universität Mainz and a Marie Curie Fellow at the Goethe-Universität Frankfurt am Main. He is coeditor, with Marcus Willaschek, Georg Mohr, and Jürgen Stolzenberg, of the *Kant-Lexikon* (forthcoming).

ANNE MARGARET BAXLEY is Associate Professor and Director of Undergraduate Studies in Philosophy at Washington University in St. Louis. Her book is on Kant's account of virtue, *Kant's Theory of Virtue: The Value of Autocracy* (Cambridge, 2010).

B. SHARON BYRD was Professor in the School of Law, Friedrich Schiller University. She coauthored (with Joachim Hruschka) *Kant's Doctrine of Right: A Commentary* (Cambridge, 2010), and authored many articles on Kant's theory of right, including "Kant's Theory of Punishment: Deterrence in Its Threat, Retribution in Its Execution," *Law and Philosophy* (1989); and "The State as a Moral Person," *Proceedings of the Eighth International Kant Congress* (1995).

LARA DENIS is Professor of Philosophy and Director of the Ethics Program at Agnes Scott College. She is author of *Moral Self-Regard: Duties to Oneself in Kant's Moral Theory* (2001) and editor of *Kant's Metaphysics of Morals: A Critical Guide* (Cambridge, 2010).

STEPHEN ENGSTROM is Professor of Philosophy at the University of Pittsburgh. He is a coeditor of *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty* (Cambridge, 1996) and the author of *The Form of Practical Knowledge: A Study of the Categorical Imperative* (2009).

PATRICK R. FRIERSON is Associate Professor of Philosophy at Whitman College. He is author of *Freedom and Anthropology in Kant's Moral*

Philosophy (Cambridge, 2003), *Kant's Questions: What Is the Human Being?* (2013), and *Kant's Empirical Psychology* (Cambridge, forthcoming) and coeditor of Kant's *Observations on the Feeling of the Beautiful and the Sublime* (Cambridge, 2011).

JEANINE GRENBERG is Professor of Philosophy at St. Olaf College in Northfield, Minnesota. She is author of *Kant and the Ethics of Humility: A Story of Dependence, Corruption and Virtue* (Cambridge, 2005), and *Kant's Defense of Common Moral Experience: A Phenomenological Account* (Cambridge, 2013).

PAUL GUYER is Jonathan Nelson Professor of Humanities and Philosophy at Brown University. He is coeditor, with Allen Wood, of the *Cambridge Edition of the Works of Immanuel Kant*. Among his recent works are *Kant's System of Nature and Freedom* (2005), *Values of Beauty: Historical Essays in Aesthetics* (Cambridge, 2005), *Kant* (2006), *Knowledge, Reason, and Taste: Kant's Response to Hume* (2008).

JOACHIM HRUSCHKA Professor in the School of Law, Friedrich Alexander University, Erlangen, Germany, is coauthor (with B. Sharon Byrd) of *Kant's Doctrine of Right: A Commentary* (Cambridge, 2010). Among his many other publications are *Das deontologische Sechseck bei Gottfried Achenwall im Jahre 1767* (1986).

MANFRED KUEHN is Professor of Philosophy at Boston University. He is author of *Scottish Common Sense in Germany* (1987) and *Kant: A Biography* (Cambridge, 2001), and coeditor, with Heiner F. Klemme, of the two volume work *Immanuel Kant: An International Anthology of Essays on Kant* (2000).

ROBERT B. LOUDEN is Distinguished Professor of Philosophy at the University of Southern Maine. His publications include *Kant's Impure Ethics* (2000) and *Kant's Human Being* (2011). Loudon is also coeditor and translator of two volumes in *The Cambridge Edition of the Works of Immanuel Kant*.

J. B. SCHNEEWIND is Professor Emeritus of Philosophy at Johns Hopkins University. He is editor of *Moral Philosophy from Montaigne to Kant* (Cambridge, 2003), as well as of the Cambridge Edition translation of Kant's *Lectures on Ethics*. He is author of *Sidgwick's Ethics and Victorian Moral Philosophy* (1977) and *The Invention of Autonomy* (Cambridge, 1998).

OLIVER SENSEN is Associate Professor and Director of Graduate Studies in Philosophy at Tulane University. He is the author of *Kant on Human Dignity* (2011), as well as the editor of *Kant on Moral Autonomy* (Cambridge, 2012).

HOUSTON SMIT is Associate Professor of Philosophy at the University of Arizona. He specializes in the history of medieval and early modern philosophy and is currently writing a book with the working title “Kant’s Theory of Cognition.”

JENS TIMMERMANN is Reader in Moral Philosophy at the University of St. Andrews. He is the author of *Sittengesetz und Freiheit* (2003) and *Kant’s Groundwork of the Metaphysics of Morals: A Commentary* (Cambridge, 2007). Among the several works of which he is an editor is *Kant’s Groundwork of the Metaphysics of Morals: A Critical Guide* (Cambridge, 2009).

MARK TIMMONS is Professor of Philosophy at the University of Arizona. He has published widely on topics in metaethics, normative ethics, and Kant’s ethics. He is editor of *Oxford Studies in Normative Ethics*, and is currently working on a book with Terry Horgan entitled *Illuminating Reasons: An Essay in Moral Phenomenology*.

ALLEN W. WOOD is Ruth Norman Halls Professor at Indiana University and Ward W. and Priscilla B. Woods Professor emeritus at Stanford University. He is author of ten books, including *Kant’s Ethical Thought* (Cambridge, 1999), *Kantian Ethics* (Cambridge, 2008) and *The Free Development of Each* (2014).

Foreword

Kant was a popular teacher until late in his long career. Even when he was a young instructor, he attracted older residents of Königsberg as well as many students to his lectures. But the lectures were designed for the students, who were usually younger than conventional undergraduates are now. Kant was helping to bring them up. He was teaching them the rudiments of morality as well as moral philosophy. Although they had to be able to read – or puzzle out – enough Latin to cope with the Baumgarten textbooks that Kant used, they were not very well educated. Students normally took the moral philosophy course as part of their first-year curriculum. It was part of the preparation for more advanced studies of law or medicine or theology. These were the professions for which students were educated.

Lecture classes met four times a week, for forty or forty-five minutes. There were additional discussion sessions. In addition to notes from the ethics courses, notes have survived from several other courses that Kant taught – for example, those on logic, anthropology, metaphysics, and physical geography. Students generally kept their notes in bound volumes, often inscribed with the student's name. The notes are conventionally identified by using the name of the student whose copy is involved. Copies of the notes circulated widely among students and others during Kant's lifetime. Many of the notes have been known to scholars for a long time. Some notes once available are now lost, perhaps destroyed during World War II. And some are fairly recent discoveries.

The first publication of notes from Kant's courses on moral philosophy was that by Menzer, in 1924 (see Menzer 1924a). His edition was the basis for the 1930 translation by Louis Infield, which was frequently reprinted. A three-volume edition of the ethics notes, edited by Gerhard Lehmann, was published in 1974 as part of the *Akademie-Ausgabe* of Kant.¹ Although the

¹ This is volume 27 of the Academy edition of Kant's works (see AA 27). It consists of three separate parts, paginated continuously.

editing has been the target of much criticism the edition has not yet been replaced and serves as the main source for work on the notes. A previously unavailable manuscript of notes on the ethics lectures, in a notebook owned by a student named Johann Friedrich Kaehler, turned up in 1997. It was edited by Werner Stark and published in 2004 (see Stark 2004a).

Stark makes strong claims for Kaehler. He admires Menzer's work and is highly critical of Lehmann's. He thinks Kaehler should replace all of these printings of the ethics notes. One reason for his assurance is that there are twelve other texts from the middle of the 1770s that are the same as Kaehler (Stark 2004b, cf. his 1999). He finds it astonishing that there should be so many identical sets from such a short period of time. Of no other lecture notes, Stark asserts, is this true. He takes this as evidence that many students found in these notes a "well worked out and very comprehensive report of Kantian moral philosophy." Moreover, in the preparation of the text, Stark tells us, modern editorial technology not available to Lehmann has been used (Stark 2004b, 392f).

Stark's claim for Kaehler is no doubt warranted. But this poses a problem for those who do not wish to read the notes in the original German. Kaehler was not available in the mid-1990s when Peter Heath translated a substantial selection of ethics notes using the Lehmann edition. Heath's translation is the fullest, and the most likely to be used by readers of English.² How reliable is it?

Stark gives some comparisons of the texts of Menzer, Collins, and Kaehler (Stark 1999, 89–97). I have myself done a comparison of the Kaehler text with Collins as in Heath's translation of it. Stark ignores some of the differences in paragraphing, single words, and short sentences or parts of sentences. Many of the discrepancies that he notes are trivial, for example, *stärkste* (*strongest*) for *höchste* (*highest*) (Stark 1999, 91). But there are a few important differences. The most substantial are certain passages in Kaehler but not in Collins. Heath has, however, used passages from notes taken by a student named Mrongovius to make up for the larger of these gaps in Collins.³ In these locations, Kaehler has essentially identical passages, thus validating Heath's additions. Another difference is also substantial. Stark argues convincingly that a dozen lines of Collins (27:375:12–24, pages 148f. of Heath's translation), should not be included. The disputed text begins, "So long as we acknowledge . . ." and ends, "defies the latter's

² In Heath and Schneewind 1997. The Collins translation provided by Heath will be referred to as "Heath-Collins."

³ Heath-Collins 68–73 is taken from 27:1425–1430; Heath-Collins 81f is taken from 27:1438.

intention.” Stark says that it is probably a copyist’s error, inserting someone’s later marginal annotation into the body of the notes (see Stark 2004a, 225, n. 179).

The other differences make – as far as I can see – no philosophical difference to what Kant said. A few examples will suffice. There are differences of paragraphing here and there. Collins has Kant attributing to Socrates the saying, “My friend, there are no friends” (Collins 27:424, Heath-Collins, 185), while Kaehler has him attributing it to Aristotle (Kaehler 295). Heath-Collins 215 differs by a few sentences from Kaehler 353f. Here and there Collins has a sentence not in Kaehler. Apart from these and perhaps a dozen other minor divergences, Collins and Kaehler are the same. The very large amount of overlap among Kaehler, Collins, and Heath-Collins should if anything strengthen our confidence in Heath-Collins as representing Kant’s lectures. In my opinion, readers who do not use Kaehler can use Collins or Heath’s accurate and fluent translation of it to get as good an idea as we can have of what Kant told his students.

The notes that students took as they attended Kant’s lectures on ethics are a rich source of information about his views on morality, the metaphysics of morals, and religion. They add general claims as well as fascinating details to what Kant says in his published writings on these topics. They enable us to trace some aspects of the development of his philosophical views of morality. They present Kantian views that are hardly if at all discussed in the published work – for instance, Kant’s understanding of the history of moral philosophy (discussed by Wood in the present volume). And they give us a glimpse of the passion with which Kant pursued certain ideas – perhaps most notably his strong opposition to any form of religion that relies on formulae, ritual, and dogma rather than on living a moral life from purely moral motives. They also show us Kant urging his students to enjoy life as fully as possible, of course within the boundaries set by the moral principle. No ascetic, he! Quite the contrary. The essays in the present volume will help the reader work through the rich and complex material in the notes. They will add to our understanding of Kant’s moral philosophy as a whole, and perhaps help us even to enjoy it more.

J. B. SCHNEEWIND

Translations and abbreviations

Unless footnoted otherwise in their essays, authors use translations from the *Cambridge Edition of the Writings of Immanuel Kant*, series editors Paul Guyer and Allen W. Wood (Cambridge University Press, 1992–). Page citations are to *Kants gesammelte Schriften*, Ausgabe der Königlich-Preußischen Akademie der Wissenschaften (Berlin: Walter de Gruyter, 1902–) using the volume:page number format or the volume:page number:line format; exceptions are noted. In some cases, section numbers or headings are used in addition to or instead of page numbers. Abbreviations that authors use within citations include the following.

- | | |
|-------|---|
| AA | Akademie-Ausgabe |
| CE | Cambridge Edition |
| ApH | <i>Anthropologie in pragmatischer Hinsicht</i> [1798] (AA 7)
<i>Anthropology from a Pragmatic Point of View</i> , Robert B. Loudon (trans.) (CE <i>Anthropology, History, and Education</i> , Robert B. Loudon and Günter Zöller [eds], 2007) |
| BBGSE | Bemerkungen zu den <i>Beobachtungen über das Gefühl des Schönen und Erhabenen</i> [1764–1765] (AA 20). Also available as Immanuel Kant, <i>Bemerkungen in den “Beobachtungen über das Gefühl des Schönen und Erhabenen,”</i> Marie Rischmüller (ed.), vol. 3 of <i>Kant-Forschungen</i> , Reinhard Brandt and Werner Stark (eds.) Hamburg: Felix Meiner Verlag, 1991
“Remarks in the <i>Observations on the Feeling of the Beautiful and the Sublime</i> ” in Immanuel Kant, <i>Observations on the Feeling of the Beautiful and Sublime and Other Writings</i> , Patrick R. Frierson and Paul Guyer (trans., eds.), part of the <i>Cambridge Texts in the History of Philosophy</i> series, Cambridge University Press, 2011
Excerpts also in CE <i>Notes and Fragments</i> , Paul Guyer (ed.), Curtis Bowman, Paul Guyer, Frederick Rauscher (trans.), 2005 |

- BGSE *Beobachtungen über das Gefühl des Schönen und Erhabenen* [1764] (AA 2)
Observations on the Feeling of the Beautiful and Sublime, Paul Guyer (trans.) (CE *Anthropology, History, and Education*, Robert B. Loudon and Günter Zöller [eds.], 2007)
- Br. Briefe (AA 10–13)
 Kant's letters (CE *Correspondence*, trans., Arnulf Zweig [ed.], 1999)
- C Moralphilosophie Collins [reflecting notes originally from 1774–1777; transcribed for or by Collins, 1784–1785] (AA 27)
 Collins notes on Kant's moral philosophy lectures (CE *Lectures on Ethics*, Peter Heath [trans., ed.] and J. B. Schneewind [ed.], 1997)
- F Naturrecht Feyerabend [1784] (AA 27) A new edition is being prepared by Heinrich P. Delfosse, Norbert Hinske, and Gianluca Sadun Bordoni. So far, one volume has appeared: *Stellenindex und Konkordanz zum "Naturrecht Feyerabend,"* Teilband I: *Einleitung des "Naturrechts Feyerabend"* (Stuttgart-Bad Canstatt: Frommann-Holzboog, 2010).
 Kant's lectures on natural right (to appear, translated by Frederick Rauscher, in CE *Lectures and Drafts on Political Philosophy*, Frederick Rauscher and Kenneth Westphal [eds.], in press)
- G *Grundlegung zur Metaphysik der Sitten* [1785] (AA 4)
Groundwork of the Metaphysics of Morals (CE *Practical Philosophy*, Mary J. Gregor [trans., ed.], 1996)
- H Praktische Philosophie Herder [1762–64] (AA 27)
 Herder notes on Kant's lectures on practical philosophy.
 Translated selections appear in CE *Lectures on Ethics*.
- HN Handschriftlichen Nachlass (AA 14–23)
 Handwritten remains (selections in CE *Notes and Fragments*, Paul Guyer [ed.], Curtis Bowman, Paul Guyer, Frederick Rauscher [trans.], 2005, and in CE *Lectures and Drafts on Political Philosophy*)
- I "Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht" [1784] (AA 8)
 "Idea for a Universal History with a Cosmopolitan Aim" (CE *Anthropology, History, and Education*)
- Kaehler Kaehler's transcription of Kant's lectures on ethics [1777]. Available in Stark 2004a.

- KpV *Kritik der praktischen Vernunft* [1788] (AA 5)
Critique of Practical Reason (CE *Practical Philosophy*)
- KrV A/B *Kritik der reinen Vernunft* [1781, 1787], cited by A (first edition) /
 B (second edition)
Critique of Pure Reason, Paul Guyer and Allen W. Wood
 (trans., eds.) (CE, 1998)
- KU *Kritik der Urteilkraft* [1790] (AA 5)
Critique of the Power of Judgment, Paul Guyer (trans., ed.)
 (CE, 2000)
- L Vorlesungen über Logik (AA 9, 24)
 Lectures on logic, J. M. Young (trans., ed.) (CE *Lectures on
 Logic*, 1992)
- M I Moral Mrongovius I [1774–1777] (AA 27)
 Mrongovius notes on Kant’s moral philosophy (brief pas-
 sages used to supplement Collins notes in CE *Lectures on Ethics*)
- M II Moral Mrongovius II [1784–1785] (AA 29)
 Mrongovius notes on Kant’s moral philosophy lectures
 (selections in CE *Lectures on Ethics*)
- MA “Mutmaßlicher Anfang der Menschengeschichte” [1786] (AA 8)
 “Conjectural Beginning of Human History” (CE
Anthropology, History, and Education)
- MS *Die Metaphysik der Sitten* [1797–1798] (AA 6)
Metaphysics of Morals (CE *Practical Philosophy*)
- P Praktische Philosophie Powalski [1782–83] (AA 27)
 Powalski notes on Kant’s practical philosophy lectures
- PS Preisschrift – “Untersuchung über die Deutlichkeit der
 Grundsätze der natürlichen Theologie und der Moral” [1764]
 (AA 2)
 “Inquiry Concerning the Distinctness of the Principles of
 Natural Theology and Morality” (CE *Theoretical Philosophy*,
 1755–1770)
- R Reflexionen (AA 14–19)
 Selections appear in CE *Notes and Fragments* and CE *Lectures
 and Drafts on Political Philosophy*. Citations include Reflexionen
 number.
- RGV *Die Religion innerhalb der Grenzen der bloßen Vernunft* [1793]
 (AA 6)
Religion within the Boundaries of Mere Reason (CE *Religion
 and Rational Theology*)

- RL *Metaphysische Anfangsgründe der Rechtslehre* [1797] (AA 6)
Metaphysical Foundations of the Doctrine of Right, part one
and the first published part of the *Metaphysics of Morals* (CE
Practical Philosophy)
- TL *Metaphysische Anfangsgründe der Tugendlehre* [1797] (AA 6)
Metaphysical Foundations of the Doctrine of Virtue, part two
and the second published part of the *Metaphysics of Morals* (CE
Practical Philosophy)
- TP “Über den Gemeinspruch: Das mag in der Theorie richtig sein,
taugt aber nicht für die Praxis” [1793] (AA 8)
“On the Common Saying: That May Be Correct in Theory,
but It Is of No Use in Practice” (CE *Practical Philosophy*)
- ÜP *Über Pädagogik* [ed. Friedrich Theodor Rink (Königsberg:
Friederich Nicolovius, 1803)] (AA 9)
Lectures on Pedagogy, trans. Robert B. Loudon (CE
Anthropology, History, and Education)
- V Die Metaphysik der Sitten Vigilantius [1793–94] (AA 27)
Vigilantius’s notes on Kant’s lectures on the metaphysics of
morals (CE *Lectures on Ethics*)
- VA Vorlesungen über Anthropologie (AA 25)
Lectures on anthropology (selections to appear in CE *Lectures
on Anthropology*, Allen W. Wood and Robert B. Loudon [eds.],
2013)
- VMS Vorarbeiten zu *Die Metaphysik der Sitten* (AA 23)
Preparatory work on the *Metaphysics of Morals*
- VpR Vorlesungen über die philosophische Religionslehre [1783–86?]
(AA 28)
Lectures on the philosophical doctrine of religion, Allen
W. Wood (trans.) (CE *Religion and Rational Theology*)
- ZeF “Zum ewigen Frieden” [1795] (AA 8)
“Toward Perpetual Peace” (CE *Practical Philosophy*)

Introduction

Lara Denis and Oliver Sensen

I The importance of Kant's lecture notes

Anyone interested in Kant's moral philosophy will find his *Lectures on Ethics* to be an invaluable resource. This is for several reasons (similarly Naragon 2010a):

First, in the *Lectures* Kant treats the same topics of his published writings more elaborately and in a more accessible manner: he uses more examples to explain his points, and uses a more colloquial language. Second, Kant also addresses topics which are not covered in his other works, and the *Lectures* help to correct a picture of Kant as a purely abstract thinker. Kant did not only talk about a priori principles, but was a keen observer of all aspects of human life; he even tells his students how to make polenta! Third, the *Lectures* help to reconstruct the development of Kant's views on moral philosophy. The different sets of notes which are preserved are from different stages of his career; reading them in comparison with one another indicates where his views have changed over time.¹ One can see Kant grapple, for instance, with the problem of what moral obligation is, and which role feelings and God play in its solution. Fourth, the notes also help to enlighten the context for Kant's moral philosophy. For more or less thirty years, from the early 1760s to 1794, Kant used the same two textbooks for his lectures on moral philosophy: Alexander Gottlieb Baumgarten's (1714–1762) *Initia philosophiae practicae primae* (*Introduction to Practical First Philosophy*), and *Ethica philosophica* (*Philosophical Ethics*).² These textbooks were the background against which Kant developed his vocabulary, questions and positions. Kant's *Lectures* therefore enlighten the context in which Kant wrote.

¹ Kant lectured from 1755–1796 (cf. Stark 2004b, 380–385). A list of when Kant gave which lecture can be found in Arnoldt 1909, 173ff, as well as Naragon 2010b.

² Cf. Lehmann 1979, 1047; Schneewind 1997a, xix–xxv; and Stark 2004b, 388f. Both of Baumgarten's works are reprinted in the Kant's Academy edition, the *Initia* with the notes Kant wrote in the margins (cf. AA 19:5–91), as well as the second and third edition of the *Ethica* (cf. AA 27:733–1015).

By contrast, there are doubts about the reliability of the lecture notes as an accurate statement of Kant's views. For one, the notes are handwritten by students. The grammar, fluency, as well as the dates by which some of them are signed suggests that they were mostly written at home after the lecture, and are not a stenographic direct recording of what Kant said.³ Furthermore, it is not always clear who wrote a particular transcript in the first place. For instance, there is a group of thirteen sets of notes – they include, among others, the notes with the names of Collins, Kaehler, Brauer, Kutzner, Mrongovius I⁴ – which are in large parts so identical that most likely one has copied the text from another set years later.⁵ Werner Stark has argued that Kaehler is the most reliable representative of Kant's lecture of this group, but even this would not mean that Kaehler was the original note taker.⁶ An additional problem is that the lecture notes contain obvious mistakes. Some words or passages clearly do not belong there. The reader who relies on an English translation might be subjected to further challenges in getting at Kant's precise meaning.⁷

Another problem is that is unclear how carefully Kant scripted his ethics lectures. Scholars still debate whether Kant spoke extemporaneously or read from a prepared booklet. One of his students, Reinhold Bernhard Jachmann, reported that Kant lectured with considerable spontaneity, often merely referring to notes he had written in the margins of the textbook,⁸ or from an outline on a small piece of paper (cf. Jachmann 1993, 116f). However, for lectures in other disciplines Kant had prepared a little booklet that contained his lectures. Although there is some indication that he might have used a prepared booklet for his lectures in moral philosophy as well, we do not have such a booklet for Kant's lectures on ethics (cf. Stark 2004b, 389–391; cf. also Lehmann 1979, 1040, note 9).⁹

³ Cf. AA 25:lxii–lxxiv; Lehmann 1979, 1050; Stark 2004b 392, 400f.

⁴ Paul Menzer used the latter three for his 1924 edition of Kant's lectures, which was later translated into English by Louis Infield.

⁵ On this see Lehmann 1979, 1041; Stark 2004b, 392–401; Menzer 1924b, 325 and Krauß 1926.

⁶ Cf. Stark 2004b, 392–401. For a brief response to Stark's claim, see Schneewind's foreword to this volume.

⁷ Jens Timmermann gives a first list of mistakes and translation problems for Mrongovius II in his contribution to the volume.

⁸ As Werner Stark indicates, it is unlikely that Kant merely read his notes in the margins during the lectures, as the amount of notes regarding Baumgarten's *Initia* (around fifty-five thousand words) is much longer than the corresponding part in the lectures (around thirty-three thousand words), cf. Stark 2004b, 397, note 80.

⁹ For an account of Kant as a lecturer, see Jachmann 1993, 116–120; Stark 1995; Schneewind 1997a; and Kuehn 2001, 105–110, 125, 129–134, 158–163, 204–218.

So by themselves the *Lectures on Ethics* are not an indubitable statement of Kant's views. But if one keeps these problems in mind, they are uniquely valuable for a complete understanding of his moral philosophy.

2 A critical guide to the *Lectures on Ethics*

The aim of the present volume is to present a Critical Guide to the notes on Kant's lectures on moral philosophy. The ambition is not to provide a full and comprehensive commentary, but to address topics and questions of strong interest to readers with expertise in Kant's ethical and political thought. The volume seeks to enrich readers' understanding of Kant's moral philosophy and its development by shedding much merited – and much needed – light on his *Lectures on Ethics*. The Guide comments on the texts as they are available in the Cambridge Edition of Kant's *Lectures on Ethics*, edited by Peter Heath and J. B. Schneewind, which contain a representative selection of the lectures available.¹⁰ The volume presents an English translation of the four sets of notes described below:

From the time before Kant became a professor in 1770, the Heath and Schneewind volume presents lecture notes by Johann Gottfried Herder. Herder studied under Kant between the summer of 1762 and November 1764. His notes on moral philosophy – lecture notes by Herder are the only ones we have from the time before 1770 (cf. Stark 2004b, 375f) – record a lecture during that period, but it is still debated when exactly that lecture occurred (cf. Schwaiger 2000, 180–183). Unlike the other set of notes, the Herder transcript was probably directly written down during Kant's class. These notes are elliptical, however, in that they seem to focus only on issues in which Herder was interested (cf. Lehmann 1979, 1046, 1048). Nonetheless, the notes provide an important window into an early stage of Kant's moral thought.

The second set of notes in the Heath and Schneewind volume were written by Georg Ludwig Collins in 1784–1785.¹¹ This is the semester in which Collins started his studies in Königsberg. Yet given the near identity with the earlier set of lectures, such as the ones by Kaehler or Brauer, the Collins notes are most likely a copy of another set. This contention is supported by the fact that the second set of Mrongovius

¹⁰ Currently one can trace about twenty-five sets of notes on Kant's lectures on moral philosophy, cf. Stark 2004b, 376–379, and Naragon 2014.

¹¹ On Collins see Lehmann 1979, 1050f. Collins and his father, a friend of Kant's, are mentioned in one of Kant's letters (cf. AA 12:179).

notes also seems to be from the same semester, but while the Mrongovius notes are close in content to the *Groundwork*, the Collins notes contain important differences.¹²

Notes by Christoph Coelestin Mrongovius are the third set translated in the Heath and Schneewind volume. Mrongovius started his studies in Königsberg in March 1781 (cf. Lehmann 1979, 1052), and there are actually two sets of notes that carry his name. The first set, Mrongovius I, is part of the group that also contain the Collins and Kaehler notes. These notes are reprinted in volume 27 of the Academy edition (cf. AA 27:1395–1581). The ones that are translated in the Heath and Schneewind volume are the second set of Mrongovius notes (cf. AA 29:597–642). These notes are of particular interest since they seem to record Kant's lectures from the winter semester 1784–1785, the time when Kant had finished writing the *Groundwork of the Metaphysics of Morals* (1785), his first major work on moral philosophy of the critical period.

Finally, the Heath and Schneewind volume contains the notes by Johann Friedrich Vigilantius in translation. The Vigilantius transcript records Kant's lectures from 1793–1794, in between the *Critique of Practical Reason* (1788), and the *Metaphysics of Morals* (1797), from the time of Kant's *Religion* (1793). The Vigilantius notes are so important because they present Kant's mature views on moral philosophy by a reliable source. Vigilantius was a government official at the time, part of Kant's social circle, his legal advisor, and present at Kant's death (cf. Lehmann 1979, 1045f). The original notes were lost at the end of World War II. The ones we have now go back to a copy of them from the nineteenth century (cf. Stark 2004b, 386 n. 46).¹³

3 Chapter overview

In order to shed light on Kant's *Lectures on Ethics*, the present volume has three parts: The first part contains one essay on Kant's relation to Baumgarten's philosophy, as well as one chapter each on the different sets of notes translated in the Heath and Schneewind volume: Herder, Collins, Mrongovius II, and Vigilantius. The chapters of the second part comment on topics of Kant's introductory sections of the *Lectures* as well as the parts that corresponds to Baumgarten's *Initia*, while the contributions to the

¹² On this, see the contributions by Timmermann and Kuehn in this volume.

¹³ Another distinct set of lecture notes that is not in the Heath and Schneewind volume are the Powalski notes. These notes are somewhat cryptic, and hard to date. They might be written down near the end of the 1770s, but their content might date from before the Collins and Kaehler notes. On the Powalski notes, see Lehmann 1979, 1043f, and Schwaiger 2000, 185–188.

third part comment on topics Kant discusses in response to Baumgarten's *Ethica*.

Chapter 1, “Kant’s Lectures on Ethics and Baumgarten’s Moral Philosophy” demonstrates how Kant’s ethical thought, as reflected in his lectures, responds to Baumgarten’s works on moral philosophy. Stefano Bacin argues that Kant chose Baumgarten’s textbooks for his classes for genuinely philosophical reasons. Kant’s thorough discussion of Baumgarten’s views, through a sort of critical dialogue, provided Kant with important clues for developing an original position, even if mostly in opposition to Baumgarten. Bacin illustrates this complex role of Kant’s “author” with a few significant examples. These examples serve also to highlight some original aspects of Baumgarten’s position in comparison to Wolff’s, Bacin argues: First, Baumgarten’s focus on obligation – and his treatment of perfection and virtue as secondary concepts – highlights what Kant holds to be the crucial problem in moral philosophy. Kant does not, however, regard Baumgarten’s account of obligation as providing a satisfying solution. Second, Baumgarten’s sharply theistic foundation of morality is rejected by Kant. Indeed, according to Bacin, in the sections of the lectures where Kant, apparently not commenting on Baumgarten, presents first elements of his own position on the moral law, he is actually giving a critical discussion of Baumgarten’s theistic thesis. Third, Kant rejects several significant aspects of Baumgarten’s division of ethical duties, thereby revealing profound differences between his conceptions of morality and Baumgarten’s.

Chapter 2, “Herder: Religion and Moral Motivation,” is especially concerned with the Herder notes and Kant’s contemporaneous ethical thought. Patrick Frierson provides an account of religion’s role in Kant’s early ethics, particularly the roles it plays in moral motivation. According to this reading, Kant insists already in the 1760s that religion cannot provide the primary moral motive, long before developing his doctrine of “respect for the moral law” as *the* moral motive. But during this period, religious motives play roles in genuinely moral motivation. Frierson argues that on Kant’s view religion adds *direct* quasi-moral motives, enriching virtue with piety and thereby completing one’s moral perfection. On this account, religion also provides very important “*mediately* motivating grounds” that are “preparatory to ethics” (H 27:14). It can humble sensuous motives (particularly those born of luxury) through resignation to the divine will, and it can alleviate moral despair (born of moral weakness) with the assurance that God can aid our moral endeavors. There are dangers of religion, Frierson notes, including the tendency toward religious laziness that leads to moral complacency,

religious fanaticisms that replace genuine morality with specious demands, and an overly speculative religion that undermines religion's practical benefits. According to this reading, however, Kant in the Herder lectures sees religion as an important part of moral life.

Manfred Kuehn's [Chapter 3](#), "Collins: Kant's Proto-Critical Position," examines the Collins lecture notes in order to ascertain how Kant's moral thought, as reflected in these notes, is related to Baumgarten's textbooks, Kant's other published lectures, Kant's reflections, and – most important – Kant's mature work. Essentially, Kuehn seeks to show how the Collins lecture notes shed light on Kant's philosophical development, as well as his moral philosophy in general. Kuehn begins by discussing the provenance of the Collins lecture notes, comparing this set of notes with others from roughly the same period, and presenting some controversies concerning the editing and publishing of the notes. Observing that Kant divided the lectures represented by the Collins notes sharply between the part based on Baumgarten's *Initia* and the part based on Baumgarten's *Ethica*, Kuehn divides his philosophical exploration of these notes along similar lines. He first considers "ethics," with a view to identifying some general features of Kant's ethical outlook and discerning how much it changed over the years. He subsequently turns his attention to "universal practical philosophy" and its relation to the categorical imperative. Although Kuehn is cautious about this, he suggests that the part of the Collins lecture notes dealing with "universal practical philosophy" corresponds roughly to what Kant later tackles in the *Groundwork of the Metaphysics of Morals* and the *Critique of Practical Reason*, and the part treating "ethics" proper corresponds roughly to the subject matter of the *Doctrine of Virtue*. Thus, Kuehn finds much to elucidate the development and mature form of Kant's foundational moral thought in the *Initia*-based part of the lecture, and much to elucidate the development and mature form of Kant's ethics, narrowly construed, in the *Ethica*-based part of the lecture.

In [Chapter 4](#), "Mrongovius II: A Supplement to the *Groundwork of the Metaphysics of Morals*," Jens Timmermann shows several ways in which reading the Moral Mrongovius II lecture notes can help us understand the ethical theory of the *Groundwork*. Moral Mrongovius II is a record of lectures on moral philosophy Kant gave in the winter of 1784–1785, when the *Groundwork* was being prepared for publication by the publisher. Some themes – for example, the good will, moral worth, hypothetical and categorical imperatives – closely mirror Kant's discussion in the published book. Other themes – such as Kant's thoughts on ancient philosophy, the highest good, the philosophy of law, and the nature of punishment – resurface

only much later, if they resurface at all. Also, there are cases in which Mrongovius II sheds light on controversial topics in Kantian ethics, for instance the workings of the categorical imperative as a formal principle and the reassuring role of religion. Through a careful analysis, Timmermann elucidates these topics.

Robert Louden's [Chapter 5](#), "Vigilantius: Morality for Humans," contends that Vigilantius's *Notes on the Lectures of Mr. Kant on the Metaphysics of Morals* deserve a special place in Kant's presentations of his ethical theory. In the Vigilantius notes, Louden argues, Kant employs a conception of the metaphysics of morals that takes human beings as its proper object. While this human perspective is also detectable in other mature works such as the *Metaphysics of Morals* (1797) and *Religion within the Boundaries of Mere Reason* (1793), it is etched much more sharply in Vigilantius. After presenting an opening argument for the importance of Vigilantius, Louden proceeds to a very brief comparison of the structures of Vigilantius and the later *Metaphysics of Morals*, offering some hypotheses for the unexpected differences between the two texts. He then turns to a more detailed examination of some key themes that Kant discusses in the opening and concluding sections of Vigilantius. In both cases, Louden draws attention to the refreshingly empirical, impure perspective of the text, while also pointing out that (at least in the case of his argument for moral faith) Kant's anthropological orientation may have inadvertently opened up a path that he himself could not endorse.

The next five chapters, constituting [Part II](#), concern a variety of preliminary and foundational topics that Kant mainly discussed early in his courses on moral philosophy, as an introduction to the lectures as well as in relation to Baumgarten's *Initia*: the highest good, the history of ethics, obligation, permissive laws, and imputation. [Chapters 6](#) and [7](#) offer insights, grounded in the lectures, into Kant's views of previous ethical systems, particularly those of antiquity. [Chapter 8](#) analyzes Kant's conception of obligation, and its relation to freedom and the moral law, as presented in the lectures on ethics. Through careful interpretation of the Vigilantius notes, the lectures from Kant's most mature period, [Chapters 9](#) and [10](#) tackle questions that bear importantly on Kant's theory of right.

In [Chapter 6](#), "Ancient Insights in Kant's Conception of the Highest Good," Stephen Engstrom draws on student records of Kant's lectures on moral philosophy in order to show that Kant's acknowledgment in his later writings of the value of happiness, and of the importance of its inclusion in a virtuous life, expresses in its essentials an understanding of morality and happiness that was firmly in place well before the publication of the

Groundwork. In those early years, and in the *Groundwork*, too, Kant not only took the highest good to include happiness as an integral component, but also followed the ancient Greeks in making the question of what the highest good consists in the starting point of his moral philosophy; Kant even saw his own answer to that question as agreeing with the general answer given by the ancients themselves, Engstrom notes. Far from regarding morality and happiness as fundamentally opposed, Kant held then even in those early years that in the highest good the good will and virtue are internally related to happiness, constituting not only the sole condition under which happiness is good, but also the sole cause through which the good of happiness can be attained.

According to Allen W. Wood's [Chapter 7](#), "Kant's History of Ethics," Kant's basic conception of reason is historical – contrary to a common reading of Kant's philosophy. Moreover, on this reading, Kant had an interesting theory about the history of ethics, though it was presented mainly in his lectures on ethics, and not in his published writings. Kant then held that ancient ethics was an ethics of ideals, while modern ethics is an ethics of principles. In Kant's view, Wood notes, the crucial juncture in the history of ethics was Christianity, which involved the insight that human beings can never attain to the ethical ideal, but should instead govern their actions as best they can through rational principles.

In [Chapter 8](#), "Moral Obligation and Free Will," Oliver Sensen argues that the lectures reveal obligation to be one of Kant's central moral concepts. While the term "obligation" hardly appears in Kant's published writings, Sensen first demonstrates how the problem of obligation occupied Kant from his very first publication on moral philosophy, and how he saw the categorical imperative as its solution. Sensen then analyzes Kant's conception of obligation as a making necessary of an action by the moral law. While obligation has an accompanying effect on feeling, the core element of obligation is that the moral law declares an action to be morally necessary. In the next section Sensen demonstrates the significance of obligation in that Kant rules out all alternative theories to his own because they cannot account for obligation and yield heteronomy. While this leaves open the possibility that there is no morality or obligation, in his final section Sensen analyzes Kant's argument that freedom guarantees the existence of unconditional obligation.

The permissive law is a topic of intense interest and debate within the secondary literature on Kant's philosophy of right. Disagreement concerns what the permissive law is, whether there is more than one type of permissive law in Kant's works, whether Kant abandoned one meaning of

permissive law in favor of another over time, and when exactly one needs a permissive law at all. [Chapter 9](#), “The Elusive Story of Kant’s Permissive Laws,” by B. Sharon Byrd, begins with the logic of obligations and permissions and continues with Kant’s works and the three permissive laws in Kant’s writings. It examines Kant’s statements in the *Vigilantius* lecture notes and expands on that discussion to include all of Kant’s examples of permissive laws relating to legally relevant actions. Byrd argues that there are exactly three types of permissive laws and that Kant’s works contain examples of these, and no more than these, three types. Any other interpretation of a permissive law within the domain of legally relevant actions exceeds the logic of obligations and permissions, a logic with which Kant was familiar.

In [Chapter 10](#), “On the Logic of Imputation in the *Vigilantius* Lecture Notes,” Joachim Hruschka places Kant’s ideas on imputation in their historical context. Kant takes concepts from Pufendorf, Christian Wolff, Achenwall, Baumgarten, and others of his time and places them into a system. The concept of the *causa libera* (free cause) and the distinction between *imputatio facti* (imputation of the deed) and *imputatio legis* (imputation of the law) are decisive in this system. Hruschka also discusses the reasons for excluding imputation on both of these levels. The chapter concludes with the question of why Kant did not include the distinction between *imputatio facti* and *imputatio legis* in the *Doctrine of Right*.

The chapters in [Part III](#) mostly address topics in ethics proper, which Kant lectured on later in the semester, in relation to Baumgarten’s *Ethica*: the nature and basis of different classes of duties, proper self-esteem, virtue and self-mastery, love, and the devilish vices. [Chapter 11](#) shows how all classes of duties, as explicated within the *Vigilantius* notes, can be understood as derived from the concept of freedom. [Chapters 12](#) and [13](#) highlight what the lectures on ethics reveal about attitudes and attributes essential for morality generally and fulfillment of duties to oneself especially: proper self-esteem, love of honor, self-mastery, autocracy, and virtue. [Chapters 14](#) and [15](#) concern love and vices contrary to our duties of love toward other human beings as depicted in Kant’s lectures on ethics. Significantly, these last two chapters present Kant’s views of self-regarding aspects of morality as crucial to his views of its other-regarding aspects. [Chapter 14](#) elucidates Kant’s views on love of others by appeal to his views on love of self. [Chapter 15](#) explicates the devilish vices in relation to self-esteem and love of honor.

In [Chapter 11](#), “Freedom, Ends, and the Derivation of Duties in the *Vigilantius* Notes,” Paul Guyer asks whether and to what degree Kant’s *Vigilantius* description of perfect duties to oneself as “flow[ing] strictly,

unconditionally and negatively from the concept of freedom” (V 27:601) is true of all duties within Kant’s system. Kant’s statements that the greatest self-consistent use of freedom is the fundamental principle of morality (e.g., C 27:346) might lead one to think it is true of all duties. Yet other statements suggest otherwise. For instance, in the *Vigilantius* lecture notes Kant says that in the case of duties of virtue, “[a]part from the freedom of the action, there is . . . another principle present, which in itself is enlarging, in that, while freedom is restricted by the determination according to law, it is here, on the contrary, enlarged by the matter or end thereof” (V 27:543). Particularly in light of Kant’s derivation of duties of virtue in the *Metaphysics of Morals*, it would be natural to read this passage as saying that duties of virtue depend not only on freedom, but, in addition, on ends, such as the happiness of others. Guyer offers an interpretation of the preceding passage according to which it says that while negative duties to both others and oneself are duties to avoid the restriction of freedom, the principle of positive duties to others as well as to oneself is that of the expansion of freedom; thus all classes of duty can in fact be derived from the concept of freedom combined with some basic facts about the human condition that bear on the realization of freedom.

Chapter 12, “Proper Self-Esteem and Duties to Oneself,” by Lara Denis, draws on the Collins and *Vigilantius* lecture notes as well as the *Doctrine of Virtue* to explore relations between proper self-esteem – a morally correct way of valuing oneself – and duties to oneself. Denis first teases apart several interconnected referents of “*Selbstschätzung*.” She describes moral forms of each, explaining some links between them and duties to oneself, and portraying these moral forms of *Selbstschätzung* as distinct, interrelated elements of proper self-esteem. Denis then shifts her focus to duties to oneself. After sketching the accounts of duties to oneself offered in Collins, *Vigilantius*, and the *Doctrine of Virtue*, Denis shows that within all three, duties to oneself are depicted as acts through which the moral attitude of self-esteem is expressed. Finally, Denis considers the bearing of the relation between duties to oneself and proper self-esteem to the primacy of duties to oneself. She suggests that attending to this relation enhances our appreciation of the primacy of these duties. Moreover, Denis contends that we should understand Kant’s most provocative, even troubling, claims that those who violate duties to oneself make themselves worthless and contemptible as attempts to convey the primacy of duties to oneself by appealing to the self-esteem of his audience.

Kant’s theory of virtue has yet to receive the systematic interpretation that has been granted to his moral theory in his most widely read ethical

works, the *Groundwork* and the second *Critique*. Moreover, this account of virtue has not been well understood. In light of the fact that Kant tends to identify virtue with strength of will in accordance with duty in the face of contrary feelings and inclinations, readers have wondered whether Kant's conception of virtue amounts to anything other than a recipe for mere continence. In [Chapter 13](#), "Virtue, Self-Mastery, and the Autocracy of Practical Reason," Anne Margaret Baxley analyzes four key aspects of Kant's account of virtue we find in the *Lectures on Ethics*, a resource that sheds valuable light on Kant's considered views of moral character, its importance in an ethical life lived in accordance with norms of practical reason, and its place in his overall theory of value. The comprehensive picture uncovered by Baxley's analysis of the *Lectures on Ethics* reveals that Kant's theory of virtue accords with his foundational commitments in his moral theory, while at the same time representing a complex and distinctive account of moral virtue.

In [Chapter 14](#), "Love," Jeanine Grenberg articulates Kant's suspicions about love of others as a moral motive, and assesses the extent to which he nonetheless welcomes love of self and others into his duty-based ethic. According to Grenberg, Kant worries about how love of self corrupts love of others, making it illusory and useless. Although he briefly flirts with a self-loving well-being as a corrective to these excesses, ultimately, in the movement from the Herder to the Collins lectures, Kant instead solidifies a moralized response to the problem centered on an esteem for self, gained through disciplining one's sensible motives. On this reading we thus see, especially in the movement from the earlier Herder lectures to the later Collins lectures, a hint of how Kant's now-familiar duty- and respect-based (instead of love-based) ethic emerges from suspicion of the motive of love for others, rooted ultimately in suspicion of the sensible motive of love for self. Despite Kant's constraints on the introduction of love into this emerging principle-based ethics, Grenberg argues, these lectures, when compared with Kant's published works, represent a high-water mark for the extent to which Kant welcomes love of others – especially love as a commanded feeling – into his ethics.

In the final essay of the volume, [Chapter 15](#), "Love of Honor, Emulation, and the Psychology of Devilish Vices," Houston Smit and Mark Timmons develop an interpretation of Kant's views concerning the psychology of the devilish vices of envy, ingratitude, and *Schadenfreude* based on remarks in the *Lectures on Ethics*. Of particular interest is the extent to which these vices share an underlying psychological unity, a thesis suggested by Kant's remarks in the lectures where he traces their etiology to "emulation" – an

instinct implanted in human nature and manifested in self–other comparisons. When properly guided, emulation figures in the development of true love of honor, central to possessing a virtuous character. But it can also lead to the devilish vices, whose possession involves a mere craving for honor. For instance, in the *Vigilantius* notes, it is claimed that envy has its “immediate ground” in emulation. In developing their interpretation of the psychology of the devilish vices, the authors address the following questions: What is Kant’s conception of true love of honor? What role does emulation play in acquiring a true love of honor? How is emulation involved in coming to have the devilish vices? What makes envy, ingratitude, and *Schadenfreude* devilish? Is there an underlying psychological unity among these vices? If so, what exactly is the psychological profile that these vices share?

The chapters presented in this volume illustrate the richness of the *Lectures on Ethics* as a resource for those seeking a fuller, deeper understanding of Kant’s moral philosophy.

PART I

The Sources

Kant's lectures on ethics and Baumgarten's moral philosophy

Stefano Bacin

I “A good analyst”

The notes from Kant's lectures in the various philosophical disciplines cannot be considered self-standing texts, not only because what we are reading was not authored by Kant himself but jotted down by his students, and later sometimes copied and compiled in different ways, but also because Kant's remarks in the lectures are not meant as original treatments of the subject, but as comments on a handbook. A careful comparison with the textbooks is therefore required for a full understanding of the lectures.

For his entire career Kant taught moral philosophy following Baumgarten's *Initia philosophiae practicae primae* (*Elements of First Practical Philosophy*, 1760; hereafter *Initia*) and *Ethica philosophica* (*Philosophical Ethics*, 1740; hereafter *Ethica*).¹ They provide the outline for Kant's moral philosophy lectures.² In the first main part of his classes, he comments on the *Initia*, that present the foundational issues of practical philosophy, then passes on to the *Ethica*, discussing quite in detail most aspects of the doctrine of ethical duties. Thereby, Kant refers to the main division of practical philosophy characteristic of the Wolffian tradition, that is, the distinction between universal practical philosophy and ethics proper: the first one devoted to a preliminary foundational clarification of the general concepts of the moral value of actions, of law and obligation, of imputation and conscience, the second one presenting

¹ The indication that Kant would have first taught on Friedrich Christian Baumeister's *Elementa philosophiae* is probably a mistake: cf. Stark 1993, 327, n. 1; and Schwaiger 1999, 34ff. (Note that the publication dates of the *Initia* in Schneewind 1997a, xxi are incorrect.) The Latin-English translations of Baumgarten's *Initia* and *Ethica* are my own.

² See, for instance, the helpful concordance of the Kaehler notes with *Initia* and *Ethica* in Werner Stark's edition (Stark 2004a), and the comparison of the Vigilantius notes with the *Initia* in Ludwig 1988, 54–56.

an extended doctrine of ethical duties.³ When Kant refers to universal practical philosophy, in fact, he also has mostly Baumgarten in mind.⁴ Like in the other disciplines, the only parts of the lecture notes that do not refer to Baumgarten's text are the introductory sections, in which Kant usually gives a general explanation of the difference between theoretical and practical philosophy and what is the subject matter of practical philosophy, along with a summary of the main views of the "ancients" on the highest good – all themes having no correspondence in the very brief "Prolegomena" of the *Initia*.⁵ Interestingly, Kant's introductions become longer and longer with the years, coming to include a sort of anticipation of the main tenets of his own account (compare, for instance, the Powalski notes with Kaehler or Vigilantius).⁶

Because Kant and his colleagues were required by the government to follow a textbook in their teaching, however, this could raise the suspicion that for Kant Baumgarten's works were little more than a teaching prop, providing only a list of the topics to be presented to the students.⁷ Initially, Kant announced that he would lecture on them only "for the time being" (AA 2:311). Even that he never actually substituted them could suggest that their contents were not important enough to bother looking for better textbooks. This seems quite unlikely, though, once we consider some facts. First, Kant drew on Baumgarten's works not only in moral philosophy, but also in his metaphysics classes, and used a work of Baumgarten's student Georg Friedrich Meier for the logic lectures.⁸ If Kant would simply have wanted to teach on a handbook of generic Wolffian inspiration, he could have followed the choice of older philosophy teachers in Königsberg,

³ On the idea of "universal practical philosophy," see Schwaiger 2005. On the significance of the universal practical philosophy for the development of Kant's project of a metaphysics of morals, see Schwaiger 2001 and Bacin 2006.

⁴ Note that in G 4:391 he speaks of "the authors" of universal practical philosophy, in the plural.

⁵ In this respect, Kant's introductions to his lectures resemble more the introductory sections of the *Allgemeine praktische Weltweisheit* of Baumgarten's student Georg Friedrich Meier: cf. Meier 1764, §§ 1–23.

⁶ Note, by contrast, that the comparison of the lecture notes with the textbooks allows the extension of occasional *lacunae* in the notes to be assessed. This is the case, for instance, with the abrupt beginning of the Herder notes, lacking the first pages. A look at the *Initia* shows that, in spite of the strong presence of sentimentalist vocabulary and of numerous references to Hutcheson, Kant is commenting here on Baumgarten's chapter on obligation, passing on to the section on constraint in H 27:6. Thus, the notes lack the introductory section and Kant's first remarks on obligation.

⁷ This has been suggested recently by Kuehn 2011, 17.

⁸ Furthermore, Kant did not use a textbook of either Baumgarten or Meier in his natural law classes instead of Achenwall's *Ius naturae* arguably because Baumgarten's *Ius naturae* (1763) was published posthumously and unfinished, and because Meier's *Auszug aus dem Rechte der Natur* came out only in 1769. On Baumgarten's *Ius naturae* see Scattola 2008.

like the *ordinarius* for practical philosophy Carl Andreas Christiani, who used the (second volume of the) important *Institutiones philosophiae Wolffianae* by Ludwig Philipp Thümmig, recommended by Wolff himself.⁹ Baumgarten's books appear to have been a personal choice of Kant's, since other of his colleagues taught moral philosophy on different handbooks, while the only other lecturing on Baumgarten in those years appears to have been Kant's former student and *protégé* Christian Jacob Kraus.¹⁰ Second, Kant clearly considered Baumgarten's some of the best philosophical works available at that time. He regarded the *Metaphysica* as "this most useful and thorough of all the handbooks of its type" (AA 1:503), and believed the *Initia* to be "the richest in content, and perhaps his [Baumgarten's] best book" (H 27:16). These works had been more or less recently published: if the *Ethica* came out in 1740, and was already a work of established reputation as Kant began to teach on it, the *Initia*, published in 1760, was then new in print. As Kant's teaching career began, Baumgarten's works could thus justifiably be considered as state-of-the-art. (Note also that Baumgarten was only ten years older than Kant, and thirty-five years younger than Wolff.)¹¹ Third, at the end of his career Kant still refers only to Baumgarten's works to vindicate his attitude on such a delicate matter as his respect for Christian religion and the Bible in his teaching.¹² In 1796, thus, Kant still believed that the choice of his textbooks expressed his philosophical stance.

Kant must therefore have had good reasons to choose these works for his classes. The full significance of his use of them could not have been recognized, though, because Baumgarten was considered only another Wolffian philosopher. One mistaken assumption grounding this misconception was

⁹ Cf. Oberhausen and Pozzo 1999, 302, 326, 360, 388. On the Philosophical Faculty at Königsberg in Kant's times see Kuehn 2001, 74 (on Christiani).

¹⁰ Cf. Oberhausen and Pozzo 1999, 473. Others taught moral philosophy on Feder's *Lehrbuch der praktischen Philosophie* (1770), especially after the prohibition, in 1775, of using Crusius' works in the lectures (cf. Kuehn 2001, 214f), which were also employed earlier.

¹¹ For biographical information on Baumgarten see Gawlick and Kreimendahl 2011, ix–xxx.

¹² See *Conflict of the Faculties*, AA 7:7: "As a teacher of youth – that is, I take it, in my academic lectures – I never have and never could have mixed any evaluation of the Holy Scriptures and of Christianity into my lectures. The texts of Baumgarten, which are the basis of my lectures and the only thing that could be at all relevant to such a discourse, are sufficient to prove this. For, being purely philosophical, these texts do not and cannot contain a single heading referring to the Bible or to Christianity" (translation slightly modified). Compare with *Ethica, auditori benevolo*, AA 27:738.1–4: "Non equidem is sum, qui nesciat, de diuinis etiam vera dicere tam esse periculosum, vt nec defuerint publicis scriptis asseuerantes, philosophum illico se deridendum praebere, quam primum sacra tangat, et religionis exercendae mentionem iniiciat." ["I am not such a man who ignores that to assert even true things about religious issues (*divina*) is so dangerous that there is no shortage of people claiming in public texts that a philosopher causes himself to be laughed at in this respect as soon as he touches upon sacred matters and refers to the religious practice."]

that the philosophers working along the lines of Wolff's system built one homogeneous school, without philosophically significant differences from the views of the master. A second assumption is that Baumgarten's originality was limited to the new philosophical discipline of "aesthetics." That Baumgarten could be credited with such an innovation has curiously not suggested that a unitarian picture of the Wolffians was probably too simple. A careful study of the *Initia* and the *Ethica* shows, however, that Baumgarten's views exhibit significant differences from the traditional Wolffian, and that in moral philosophy he was as original as in metaphysics and in other parts of philosophy.¹³ In the following, I shall point out some of these differences.

Kant's choice of using Baumgarten's works in his lectures must thus have been based on an appreciation of their worth and their originality within the Wolffian party, as his preference for Baumgarten and Meier over other Wolffians suggests. Kant wanted to teach on some of the most advanced works which tried to develop the Leibnizian-Wolffian approach in an original way. However, as it will soon appear, Kant's attention is almost always combined with criticism. If he had good reason to follow Baumgarten, his remarks in the lectures do not often express agreement with him.¹⁴ Ultimately he does not share the basic tenets of his position. Kant's critical remarks against Baumgarten's statements during the lectures must indeed have been numerous, as appears from the notes.¹⁵ Indeed, many remarks do not even sound like objections to the reader, because Kant (or the student taking notes) does not make the contrast with the textbook explicit. For instance, a comment like: "External religion is a contradiction. All religion is within" (Kaehler 154; C 27:330; cf. Kaehler 121; C 27:308) does not appear directed against Baumgarten until we see that it addresses precisely his definition of external religion (cf. *Ethica*, § 115).¹⁶

In one of his notes in his copy of Baumgarten's *Metaphysica* Kant writes: "the man was sharp-sighted (in little things) but not far-sighted (in big ones)"; he is "a good analyst, but not an architectonical philosopher," "a

¹³ The merit of pointing out the self-standing philosophical profile of Baumgarten belongs most prominently to Clemens Schwaiger, who has highlighted the many original aspects of his thought, and especially of his practical philosophy, in numerous important publications, most of which are collected in Schwaiger 2011; cf. also Schwaiger 1999, 49ff. Yet, the significance of Baumgarten's moral philosophy apparently still needs to be stressed, since the entries on Baumgarten in two good recent reference works like the *Dictionary of Eighteenth Century German Philosophers* and the *Continuum Companion to Kant*, oddly enough, do not even mention his practical philosophy, but only his aesthetic and metaphysics.

¹⁴ Schwaiger appropriately defines Kant's relationship with Baumgarten as a "negative dependence" (Schwaiger 2011, 126).

¹⁵ See, for instance, V 27:625: "From § 201 on, Baumgarten treats of the *officia erga animam*. Professor Kant censures his plan for the following reasons."

¹⁶ I mention a further example of implicit criticism in section 3, "The Foundation of Morality."

Cyclops among metaphysicians, who was missing one eye, namely critique" (R #508I, 18:81f).¹⁷ The same remark can express Kant's appreciation of Baumgarten's moral philosophy as well. The *Initia* and the *Ethica* are also the work of a good analyst, providing a helpful inventory of concepts, but missing their real link with the activity of reason. His account requires thus to be profoundly revised. Kant's attempt at these revisions is one core aspect in the development of his ethical thought.

Because Kant closely follows Baumgarten, virtually every topic touched in the lectures could, and should, be traced back to the *Initia* and the *Ethica*. Moreover, in virtue of Kant's long and intensive use of them, the influence of Baumgarten's terminology on Kant's vocabulary is pervasive. A full account of the relationship between Kant's lectures and Baumgarten, therefore, would require a thorough commentary, which should cover also Kant's handwritten notes on his copy of the *Initia*.¹⁸ Because this is obviously not possible here, in the following I shall focus rather briefly on a few especially important points, to show how Kant understands, and draws on, Baumgarten's moral philosophy.

2 Obligation as key concept

If the textbooks must not be regarded as merely providing a list of topics to be discussed in class, the order of the topics does matter, as it reflects some substantial thesis. This is the case with the systematic outline of Baumgarten's moral philosophy, which already shows that he cannot be taken to be simply another one of Wolff's many disciples. The *Initia* stands out because of the emphasis on obligation, whose thorough examination builds the beginning of the presentation. Thereby Baumgarten departs from Wolff's pattern, whose treatment of universal practical philosophy begins, in the first, German version, with a chapter on the "fundamental rule of actions" (cf. Wolff 1720, §§ 1–71). Before Baumgarten, Thümmig had already given priority to obligation over law (cf. Thümmig 1726, §§ 18–22), but did not provide a treatment of it comparable with Baumgarten's, whose *Initia* devotes fifty-nine dense sections to the topic. Moreover, Baumgarten's examination includes, as a second step, an analysis of the concept of "constraint" (*coactio*, corresponding to *Zwang* in Kant's German), conspicuously absent both from Wolff's and Thümmig's expositions. Most important, only

¹⁷ On Kant's contrast between "cyclopes" and architectonical philosopher(s), see Ferrarin 2013.

¹⁸ See AA 19:5–317. Kant's copy of the *Ethica*, that must have contained numerous notes as well, went missing.

in Baumgarten does the concept of obligation provide the overarching idea embracing the whole practical philosophy. The *Initia* not only opens with a chapter on obligation, but all the topics of the second and last chapter are considered as *obligantia*, that is, as elements concurring to obligation. Accordingly, practical philosophy is defined as “scientia *obligationum* hominis sine fide cognoscendarum” [“science of the *obligations* of man to be known without faith”] (*Initia*, § 1), echoing the definition of philosophy in general as “scientia qualitatum in rebus sine fide cognoscendarum” [“science of the qualities to be known in the things without faith”] (cf. Baumgarten 1761, § 1; Baumgarten 1770, § 21), and the *Ethica* is analogously presented as “scientia *obligationum* hominis *internarum* in statu naturali” [“science of the *internal obligations* of man in the state of nature”] (*Ethica*, § 1).¹⁹ In regarding obligation as “the fundamental concept” (PS 2:298, cf. 2:300) of practical philosophy, Kant certainly draws also on Baumgarten, who has the merit not only of giving the concept priority, but also of showing the unity of the entire practical philosophy as a theory of obligation. Thereby Baumgarten’s views and vocabulary took a quasi-juridical overtone, following Heinrich Koehler’s *Juris naturalis Exercitationes*.²⁰ If Kant’s conception of morality has sometimes been regarded as too close to law, it is also because of his relying on Baumgarten’s vocabulary.

Baumgarten’s originality within the Wolffian camp stands out even more clearly if we consider that, focusing on obligation, he does not mention two concepts that in Wolff and in other Wolffians had pride of place, namely happiness and virtue.²¹ In both respects, his account thereby seems to take a turn that Kant must have appreciated.²² According to the outline put forward in the *Philosophia generalis*, Baumgarten did believe that the universal practical philosophy should eventually include, after a theory of laws (*nomologia*), also a theory of virtue (*aretologia*) and a theory of happiness (*eudemonologia*), along with a “universal knowledge of man” (*anthropogno-sia universalis*).²³ No trace of such a plan was to be found in the published works on moral philosophy, though, where the concept of obligation was clearly predominant. Whereas both Wolff and earlier Wolffians such as Thümmig and Gottsched gave virtue some relevance, considering it

¹⁹ Unfortunately, the reconstruction of the Wolffian views on obligation in Hartung 1998, 148ff, does not take Baumgarten into account.

²⁰ On that, see Schwaiger 2011, 116ff, 130, and Aichele 2005.

²¹ Both are dealt with in the first chapter of the *German Ethics*: Wolff 1720, §§ 52ff and § 68ff.

²² Ameriks 2012, 64 suggests that these points (priority of obligation, imperativism, noneudaimonism) would be enough to understand why Kant chose Baumgarten’s works for his classes.

²³ Baumgarten 1770, § 149:69. Cf. Schwaiger 2011, 131f.

“readiness [*Fertigkeit*] to direct one's actions according to the law of nature” (Wolff 1720, § 64), the concept simply does not play a role in the *Initia*. The *Ethica* mentions it immediately in § 1, giving (only in the 1763 edition) “doctrine of virtue” as German equivalent for “ethics,” but does not deal with virtues in any depth in the following.²⁴ Baumgarten did probably share Wolff's basic idea, but, in absence of explicit formulations, what remains is an equation of ethics and doctrine of virtue. The title of the second part of the *Metaphysics of Morals* goes back to this remark of Baumgarten's,²⁵ and to the idea that virtue is the internal disposition of agents complying with moral demands (cf. P 27:162f; V 27:631). So understood, a doctrine of virtue clearly stands in no opposition with a doctrine of duties.²⁶ Kant elaborates on this view when he observes that “virtue” stands primarily for the moral dimension open to finite agents, who cannot aspire to sanctity, since “virtue” is the moral dimension of agents constitutively put under obligation and constraint, who have to act in spite of opposite impulses (cf. e.g. H 27:13, P 27:165 and R #6993, 19:222).

The irrelevance of happiness in Baumgarten's exposition represents an even more conspicuous departure from the Wolffian pattern (cf. Schwaiger 2011, 152). While Wolff regarded happiness as the primary goal of practical philosophy (“*finis ethicae est felicitas hominis*”),²⁷ Baumgarten hardly mentions it in the *Initia*,²⁸ and only briefly in the *Ethica* (cf. §§ 10, 13). Unlike Wolff and other Wolffians, he carefully avoids confusing perfection with happiness, probably also to overcome Joachim Lange's accusation of hedonism against Wolff (cf. Schwaiger 2011, 163ff). Baumgarten not only focuses on perfection at the expense of happiness, but also tries to achieve a more precise understanding of what perfection as normative concept should mean. He differentiates therefore between perfection as a means and perfection as an end (cf. *Initia*, § 43; *Ethica*, § 10). Kant clearly appreciates Baumgarten's rectification of Wolffian perfectionism, especially the distinction between perfection and happiness.²⁹ Thus, if Kant rejects “a perfection which is in turn identical with human happiness” (Guyer 2011, 205), this

²⁴ See the quick remarks in *Ethica*, §§ 317 and 370.

²⁵ Pace Merle 1998, according to whom the term “*Tugendlehre*” was introduced by Kant.

²⁶ See also the preface to the first edition of the *Ethica* (*Auditori benevolo*), in AA 27:737.24f: “*triplex officiorum genus, cum suis quodlibet virtutibus*” [“the threefold class of duties, each of them with the corresponding virtues”].

²⁷ Wolff 1750, § 8; cf. Wolff 1720, § 45.

²⁸ Baumgarten differentiates only in *Initia*, § 98, between “internal and external happiness.” On Baumgarten's view on happiness, see Schwaiger 2011, 106ff and 152.

²⁹ See e.g. R #6487, 19:24, where Kant annotates to “*quaere perfectionem*” (*Initia*, § 43): “not merely *felicitas*.”

appears to be something that he borrows from Baumgarten's version of perfectionism. It is not Kant but Baumgarten who first untied the link between perfection and happiness that lies at the core of "the perfectionist tradition from Aristotle to the Wolffians" (Guyer 2011, 213).

Baumgarten's careful restatement of perfectionism thereby suggested the possibility of developing a noneudaimonist moral philosophy. In fact, Kant never addresses Lange's much discussed point as a possible objection to a perfectionist view, arguably because through Baumgarten he has access to a more advanced state of the debate, which does not require further discussion of the classical eudaimonist reading of perfectionism. The issue at stake now is whether the concept of perfection might be made clear enough through some distinction like Baumgarten's, and more generally, whether the concept of perfection can provide an effective criterion of moral choice at all, and for all of the ethical obligations. Kant is willing to accept "*perfice te*" as an ethical principle, provided that the meaning of moral perfection is made clear, since Baumgarten, like every Wolffian, had neglected to do that properly.³⁰ However, if the general separation between perfection and happiness was Baumgarten's achievement, Kant's own development brings him to a further restriction of the moral significance of perfection. He reaches quite early the conviction that, if made less vague, the concept of perfection can apply only to the duties to ourselves. A few notes to § 43 of the *Initia*, where Baumgarten states the principle "*quaere perfectionem*," clearly expresses the direction of Kant's reflection on Baumgarten's distinctions: first, Kant remarks "Seek perfection (*bonitas*), not agreeableness [*annehmlichkeit*]," then "the perfection of the man and the perfection of the condition are to be distinguished," but the outcome is "One cannot say that the supreme moral rule has regard to the *perfection* of others" (AA 19:125). At the same time, Kant remarks that the obligations to others refer to their happiness.³¹ The coordination of the two objective ends stated later in the *Metaphysics of Morals* can thus also be tracked back to Kant's elaboration on Baumgarten's greater care in separating perfection and happiness and in defining the proper meaning of both.

³⁰ Cf. H 27:16.27f; Kaehler 43. See, e.g., R #7268, 19:298: "The proposition 'make yourself perfect' can be seen as the principle of ethics if it is taken to say simply [*wenn er so viel sagen soll als*]: 'be good, make yourself worthy of happiness, be a good man, not a merely happy one'."

³¹ Cf. Kaehler 77; C 27:282; and M I 27:1432: "In ethics the laws have a relation to the happiness of others . . . *Ethice obligans respectu aliorum est felicitas aliorum, juridice obligans respectu aliorum est arbitrium aliorum*." Note that, as the comparison with Mrongovius and Kaehler shows, Collins' notes (and the English translation accordingly) are deceptively incomplete and miss precisely Kant's point, conflicting with the Latin sentence.

However, the necessity of a more precise idea of perfection following Baumgarten's example ultimately shows that, if we want to cover through it the entire field of morality, that concept remains nonetheless vague and is therefore not tenable as a general principle.³² Analogously, the other two main commands that, according to Baumgarten, should express the fundamental moral obligation are as much as inadequate, in Kant's eyes. Both "do good" (*fac bonum*) and "live according to nature" (*vive secundum naturam*) cannot serve as moral principles. The mere fact that Baumgarten states more than one principle exhibits a severe weakness of his account, for Kant, since "where there are already many principles in ethics, there is certainly none, for there can be only one true principle" (C 27:266; Kaehler 44). Moreover, they are practically unhelpful, as they are mere tautologies expressing empty commands that do not determine *what* to do.³³ Thus, they do not even fulfill the conditions for obligation explained by Baumgarten himself, since tautological principles cannot provide motives. They merely express the imperatival modality of obligation.³⁴

The connection of an action with motives was indeed the point of Wolff's and Baumgarten's accounts of obligation. Elaborating on Wolff's view, Baumgarten describes obligation as connecting a possible action with *causae impulsivae potiores*, with overriding impulsive causes (cf. *Initia*, §§ 12–16).³⁵ Here, however, Kant is closer to Wolff's formulation. As Kant rephrases it, obligation "is, as it were, the result of the motives" (P 27:114, cf. 27:126). He thereby rejects Baumgarten's inclusion of sensible stimuli as acceptable grounds of obligation. Quite to the contrary, he refers to motives to exclude a determination to act on mere impulses, so that the distinction between acting from stimuli and acting on motives becomes a fundamental difference in his view of action (cf. P 27:111, 122f). The Wolffian-Baumgartenian take on obligation highlights nonetheless the crucial link with an internal determination of the agent's will, in contrast with command-based accounts. Accordingly, one of the most unfortunate weaknesses of universal practical philosophy is the lack of a real clarification of the motives and their difference. This ambivalent appreciation of Baumgarten's

³² See, e.g., the distinction between perfection and moral goodness in Kaehler 43; C 27:265f.

³³ Cf. P 27:129f; Kaehler 41ff; C 27:264; V 27:517. Against tautological moral principles in general cf. e.g. *Philosophische Enzyklopädie*, 29:8.32–38.

³⁴ Note that Kant finds in Baumgarten also the concept of "imperative": cf. *Initia*, § 39. Cf. Schwaiger 1999, 165ff.

³⁵ Cf. Schwaiger 2011, 120. Baumgarten's caveat against a "chimerical ethics" (on Kant's interest in it, see Thorndike 2008 and Dyck 2012) is a particular case of his general insistence on the necessity of linking obligations with subjectively possible motivation (cf. *Initia*, § 27).

view on obligation explains why, if Baumgarten helped Kant to recognize it as “the fundamental concept,” Kant at the same time considers it “yet [so little] known” (PS 2:298, cf. 300).³⁶

3 The foundation of morality

Except for the introductory sections, Kant’s lectures follow Baumgarten’s order, so that every part of the lecture notes roughly corresponds to a few sections of the handbooks. This appears not to be the case, though, with a crucial section, present, in longer or shorter form, in all lectures. Between some comment on prohibitive and permissive laws referring to *Initia*, § 68 and brief remarks on the difference between letter and spirit of the law regarding § 74 (or, sometimes, on universal and particular laws in §§ 72f), we encounter a (usually rather conspicuous) section where Kant does not seem to present Baumgarten’s view in any way. Instead of remarks on one or another definition, we find here, often under the heading “Of the supreme principle of morality,” a critique of inadequate views on the foundation of morality that becomes a sort of description *per oppositionem* of the requisites of its principle.³⁷ These remarks, thus, appear to be, along with the introductory sections, the main genuine addition to Baumgarten in the lectures.³⁸

In spite of appearances, however, those pages also refer to Baumgarten. Since here Kant reaches the most significant point of disagreement with him, he does not limit himself to pointing out the weaknesses of his definitions, but addresses the general issue of the origin of moral laws (cf., e.g., the formulation in P 27:135.21).³⁹ In §§ 69–71 of the *Initia*, Baumgarten maintains that the natural law grounding moral obligations is, at the same time, a divine law.⁴⁰ For him, there is simply no difference between the natural law and God’s will about the free determinations of the

³⁶ Cf. Schwaiger 1999, 39f, 43ff.

³⁷ Cf. H 27:9ff; P 27:135ff (“Of the moral law”); C 27:274–8; Kaehler 55–73; M II 29:620.38–629.

³⁸ This has been suggested by Reich 2001, 387.

³⁹ Note that, as Reich observes (*ibid.*), the title “Of the supreme principle of morality,” that does not stem from the *Initia*, anticipates Kant’s formulation of the central aim of the *Groundwork*: “the identification and corroboration of the supreme principle of morality” (G 4:392).

⁴⁰ A similar view was already held by Koehler: Cf. Koehler 1738, § 330: “Deus per naturam humanam certa motiva connexit cum actibus hominum liberis, adeoque per eandem nos obligat ad actiones per se bonas exequendas. Obligatio proinde naturalis est etiam divina. Hinc *vox naturae vox Dei audit.*” [“God has linked definite motives with free human acts through human nature, so that through nature He obliges us to carry out actions that are good in themselves. Thus natural obligation is also divine. Hence *the voice of nature is the voice of God.*”] On Koehler’s equation of the “voice of nature” with God’s voice (see again § 359) and its relation to Baumgarten’s view, see Aichele 2004.

human will: "A lege naturali ad voluntatem dei circa liberas hominum determinationes, et a voluntate dei circa liberas hominum determinationes ad legem naturalem valet consequentia" ["From the natural law one can infer God's will regarding the free determinations of human beings, and from God's will regarding the free determinations of human beings one can infer the natural law"] (*Initia*, § 69). In this important respect, Baumgarten's view differs from the classical version of moral rationalism, in holding that God's will, and not simply his reason, is to be considered as the ultimate ground of the moral laws.

Baumgarten does not accept the "impossible hypothesis" of God's nonexistence put forward by Grotius (and by Gregory of Rimini before him), even though it had been endorsed also by Leibniz and Wolff: "*Neque tamen hoc posito admittitur*: 1) ius naturae late dictum s. philosophia practica esset, existeretve, etiam si non daretur deus, 2) prorsus est independens a deo, 3) ex voluntate dei nulla ratione omnino derivari potest, 4) aequè bene cognosci potest ab atheo, ac ab agnoscente divinam existentiam" ["Granted all that, it is nevertheless *not accepted*: 1) that natural law in the broad sense, that is, practical philosophy, would be or exist even if there were no God, 2) that it is thus independent from God, 3) that it cannot be derived from God's will through any reason whatsoever, 4) that it can be known by an atheist as much as by someone who recognizes God's existence"] (*Initia*, § 71). This position contrasts with Wolff's, who maintained that, "because this rule is a law because it obligates, and the obligation comes from nature, the law of nature is validated by nature itself and would hold even if man had no superior who could obligate him to it. In fact it would hold even if there were no God" (Wolff 1720, § 20; cf. § 24 and § 38).⁴¹ On the contrary, Baumgarten's view entails that, while an atheist can recognize the natural law as to its content, he cannot reach the same grasp of its full meaning that is open to the believer, as Baumgarten quite sharply formulates: "ius naturae athei s. philosophia practica, quam in suo errore perseverans cognoscere potest, destituitur ea 1) latitudine et copia, 2) dignitate materiae, 3) veritate, 4) luce, 5) certitudine, 6) vita, cuius capax est ius naturae late dictum s. philosophia practica existentiam divinam admittentis" ["the natural law of the atheist, that is, the practical philosophy that he who perseveres in error can know, is deprived of 1) the breadth and abundance, 2) the dignity of the subject matter, 3) the truth, 4) the light, 5) the certainty, 6) the life, of which the natural law in the

⁴¹ For Wolff's relation to Grotius, see also, e.g., Wolff 1718, II.viii.

broad sense, that is, the practical philosophy, of he who admits God's existence is capable"] (*Initia*, § 71). When he mentions the hypothesis "etiamsi non daretur deus" ["even if there were no God"], Baumgarten refers to the bold assertion of § 824 of the *Metaphysica*: "Si deus non actualis esset, falsum esset principium contradictionis." ["If God were not actual, then the principle of contradiction . . . would be false."⁴²] Later on, Baumgarten accordingly explains that God is both author and lawgiver of the moral laws: "deus est auctor naturae universae . . . , et omnium inde evenientium realium . . . , obligationes autem naturales sunt reale quid et positivum . . . et in eadem rationem sufficientem habent. Ergo deus est auctor obligationum, adeoque et legum naturalium" ["God is author of the whole nature . . . , and of all the realities springing from it . . . , but natural obligations are something real, and also something positive, . . . and they have in that (i.e., in nature as a whole) sufficient ground. Hence God is the author of the obligations as much as of the natural laws."] (*Initia*, §100; cf. Meier 1764, § 140). On Baumgarten's view, thus, practical philosophy is not a self-standing discipline, not only because, like every part of philosophy, it borrows concepts from metaphysics,⁴³ but also for the more important reason that practical philosophy cannot embrace its own foundation, since its ground is provided only within natural theology.⁴⁴ Thereby, Baumgarten adds to the core of Wolff's ethics a particular accentuation on God's role in the foundation of morality, probably also motivated by the project of bringing closer Wolffian philosophy and Lutheran orthodoxy, under the influence of his elder brother Siegmund Jacob Baumgarten, not only one of the most prominent German theologians at that time, but also the one who introduced Wolff's philosophy to Alexander Gottlieb.⁴⁵

The various lectures notes show that Kant never accepted this position. When he observes that in his times many moral philosophers endorse "the divine principle" (cf. Kaehler 61; M II 29:622),⁴⁶ he does not only mean the traditional divine command theorists, but also Baumgarten, who ultimately refers the origin of morality back to the creator of the universe. This fundamental disagreement with him becomes explicit only in the

⁴² Baumgarten 1739 as translated by Fugate and Hymers 2013, 285.

⁴³ Cf. *Initia*, § 87: "sola metaphysica habet principium obiectivum absolute primum domesticum" ["only metaphysics has an absolutely first objective principle of its own (*domesticum*)"].

⁴⁴ On the dependence of practical philosophy and natural law on natural theology, in Baumgarten's account, see Scattola 2008, 24f.

⁴⁵ Cf. Meier 1763, 12ff. On Siegmund Jacob Baumgarten, see Sorkin 2003.

⁴⁶ Similarly, in KpV 5:40 Kant attributes a theological account to "Crusius and other theological moralists." For a recent discussion of Kant's rejection of divine command theories, see Stern 2012, 53ff.

Vigilantius lecture: "Crusius found this necessitating person in God, and *Baumgarten likewise in the divine will*, albeit known through reason, and not positively, and on this principle a particular moral system has been erected" (V 27:510). However, Kant has always denied that morality and its laws can be referred back to an act of creation. In his eyes, Baumgarten's explanation of the foundation of morality does not build a satisfying alternative to traditional voluntarist accounts and amounts to a "theological morality, namely, a morality in which the concept of obligation presupposes the concept of God," the main weakness of such a view being that it "has no principle; or if it does have one, this is nothing but the fact that the will of God has been revealed and discovered" (AA 28:1002f). Drawing on a standard objection against voluntarism, Kant argues that, to attribute to God justice or goodness, we have to presuppose the validity and meaning of moral predicates. Therefore, he maintains that the moral laws are not derived from God's will, but the other way around: they represent our only access to God (cf. e.g. P 27:136). While Baumgarten ultimately has to refer back to natural theology to ground morality, Kant maintains that "morality . . . must not be grounded on theology, but must have in itself the principle which is to be the ground of our good conduct" (AA 28:1003) and that "religion is nothing but morality that is applied to theology" (P 27:169).

Against that view, Kant elaborates on three points found in Baumgarten himself. First, he reaffirms the idea of objective morality. A remark like "If we divide morality into objective and subjective, that is utterly absurd" is a further implicit criticism of Baumgarten, who in § 36 of the *Initia* makes precisely that distinction.⁴⁷ Kant adheres to a stronger understanding of objective morality: "for all morality is objective, and only the condition for applying it can be subjective" (Kaehler 41; cf. C 27:264). Second, he draws on the Wolffian-Baumgartenian idea that obligation depends on motives to argue that its ground cannot be external, like God's will, but must be internal to the will of the agent (cf. Kaehler 37ff; C 27:262). Kant thus takes Baumgarten's view on the foundation of morality to be in contrast with his account of obligation in terms of internal constraint, and sets the goal of combining this thought with an account of the self-standing objectivity of morality. Third, Kant borrows Baumgarten's distinction between author of the law and lawgiver as author of the obligation to distinguish, against Baumgarten, a role that God can play from one that

⁴⁷ Baumgarten probably borrows the distinction between objective and subjective morality from Koehler 1738, § 327. Compare with the passages on objective morality in Wolff listed in Schwaiger 2011, 149, n. 457.

he cannot play.⁴⁸ For Kant, God cannot be regarded as the author of the moral laws, because they have no author at all (cf. P 27:145). Along the lines of classical rationalism, God does not originate them, “just as God is no originator of the fact that a triangle has three corners” (C 27:283; Kaehler 79; cf. P 27:137; M II 29:634; V 27:547). Against Baumgarten, moreover, Kant understands the distinction between natural and positive laws to be exclusive.⁴⁹ If God cannot be the author of the moral laws, Kant suggests, at least until 1785, that God can be nonetheless regarded as the author of moral obligation (cf. H 27:10; P 27:146; C 27:277f, 283; Kaehler 61f, 79). Interestingly, along with the difference between authorship of the law and lawgiving, Kant’s distinction between principle of evaluation (*principium diiudicationis*) and principle of execution (*principium executionis*) surfaces in the lectures (cf. Kaehler 55f, 62; C 27:274, 277f). The former corresponds to the cognition of the content of moral demands as natural laws (to the “objective morality”), that has to be possible without any further assumption (cf. H 27:9.20f). The latter regards the drive to act according to those demands, for which it is necessary to assume a legislator who enforces the moral laws. Accordingly, Kant sometimes observes that acting from duty, and not from coercion, is what God demands (cf. Kaehler 51f, 79; C 27:272, 283).⁵⁰

The story of the development of Kant’s views on the foundation of morality, on the relationship between moral judgment and motivation and between morality and religion, cannot be told here, even if only with regard to the lectures.⁵¹ What is relevant, here, is that this development begins when Kant departs from Baumgarten’s position, not simply rejecting all of his results, but mostly elaborating on materials that he finds in

⁴⁸ On this distinction see Kain 2004. On Kant’s use of the author/lawgiver distinction, see also Reath 2006, 145f–16; Irwin 2004, 151ff; Irwin 2009, 156f.

⁴⁹ Cf. M II, 29:635: “Our author says: moral laws can be viewed as positive divine laws, but that is wrong, for positive laws are merely pragmatic, mere *statuta* (statutes).” Kant is referring to *Initia*, § 66: “Lex tamen et ius positiva, tam divina, quam humana, possunt simul esse naturalia, si et quatenus possunt simul ex natura actionis agentisque cognosci, sicut lex et ius naturalia, possunt etiam positiva esse, tum divina, tum humana, si et quatenus eadem ex arbitrio dei hominumve libero sufficienter cognosci possunt.” [“A positive law and a positive right (*ius*), both divine and human, can at the same time be natural, if, and insofar as, they can at the same time be known from the nature of the action and of the agent, just as a natural law and a natural right (*ius*) can also be positive, both divine and human, if, and insofar as, they can be adequately known from the free power of choice (*arbitrium*) of God or of the human beings.”]

⁵⁰ I am not implying, however, that this should be considered the only source of Kant’s distinction between *diiudicatio* and *executio bonitatis*. See Schwaiger 1999, 92ff and Schwaiger 2011, 127 for another suggestion.

⁵¹ See the relevant chapters in this volume.

Baumgarten himself,⁵² as is the case even in the one section of the lectures that would appear not to comment on him.⁵³

4 The structure of the system of ethical obligations

Compared to Kant's critical discussion of Baumgarten's foundations of practical philosophy in the *Initia*, on which I have focused so far, the parts of the lectures commenting on the *Ethica* are less innovative, as they show a more pronounced continuity in the development of Kant's thought. Even substantial parts of the *Metaphysics of Morals* are rather close to earlier treatments of the same topics in the lectures (cf. Kuehn 2011, 18f). This does not reflect Kant's attitude toward Baumgarten's account of duties, but rather Kant's primary focus on foundational issues as his ethical thought developed. More importantly, this must not be taken to entail either that Kant's views on specific ethical duties did not evolve at all, or that he limits himself to giving a more detailed exposition of Baumgarten's views, like Meier did in his *Philosophische Sittenlehre*. On the contrary, Kant devotes the same critical attention to the structure and the contents of the doctrine of ethical duties. That ethics was understood as a doctrine of duties, divided into duties to oneself, to others, and to God, cannot surprise, since both Baumgarten and Kant here follow a pattern influentially advocated by Pufendorf and mostly accepted at that time.⁵⁴ They employed the pattern differently, though, and held different opinions on the division and ranking of obligations. On this, Baumgarten had an original position, in comparison to Wolff and other Wolffians, but Kant rejects significant aspects of it. I shall briefly mention two relevant examples.

The most apparent feature of Baumgarten's *Ethica* is the priority that Baumgarten, unlike Wolff, gives to the duties to God, and the lengthy exposition devoted to them (§§ 11–149).⁵⁵ Wolff had instead maintained that duties to oneself have a priority and override the others in case of conflicting obligations (cf. Wolff 1718, II.vi, § 26 ff). This difference is

⁵² See also Kant's notes on *Initia*, § 71 in AA 19:150.

⁵³ That is, "Of the supreme principle of morality." Cf. H 27:9ff; P 27:135ff ("Of the moral law"); C 27:274–8; Kaehler 55–73; M II 29:620.38–629.

⁵⁴ I therefore do not agree with Manfred Kuehn's suggestion that the tripartite division of ethical duties "had a lasting effect on Kant" and would be the most clear example of Baumgarten's influence on him, so that "even the subdivisions of these two broad divisions are clearly indebted to the *Ethica*" (Kuehn 2011, 17f). Analogously, that "Kant would follow Wolff" in the division of duties (Guyer 2011, 200) must not be mistaken for a direct influence.

⁵⁵ See Schwaiger 2011, 140 and 154 n. Cf. Meier 1763, 35: "Er hat die Pflichten gegen Gott viel besser abgehandelt, als andere."

notably the main point for which Baumgarten feels he owes an explanation in the preface of the *Ethica*. In spite of his intellectual debt to Wolff, the duties to God had to be placed before the others, he argues, simply because they “contain the most sacred bonds of the other obligations [*quia reliquarum obligationum augustissima vincula continent*]”⁵⁶ and make it easier to comply with them (cf. *Ethica*, § 21). Baumgarten is committed to such a position because of his view on the foundation of morality. In fact, regarding the duties to God as most fundamental and primary is a characteristic feature of the expositions of authors holding voluntarist accounts.⁵⁷ If the moral demands ultimately depend on God’s lawgiving, the connection to God must be the first constitutive element in the structure of morality. This is precisely Baumgarten’s position, and applies, on his view, even if the obligations presented in the *Ethica* are to be known *sine fide*, without assuming faith. Here his intention of bringing Wolffian ethics closer to religious orthodoxy and pietism surfaces again, and Kant’s rejection of the aspects of Baumgarten’s moral philosophy most closely tied with this project shows that he did not share this aim.⁵⁸

This un-Wolffian feature of Baumgarten’s *Ethica* explains why Kant’s lectures begin the exposition of ethics proper with a longer part on religion, after Kant had already discussed similar issues in the first main part, commenting on Baumgarten’s sections on God’s authorship of moral laws.⁵⁹ However, Kant follows Baumgarten’s order, but rejects the underlying idea.⁶⁰ In the lectures of the mid-1770s, he even begins the comment on this part with another implicit, but very clear, criticism of his author: “natural religion should properly furnish the conclusion to ethics, and set the seal on morality” (Kaehler 115; cf. C 27:305), instead of being considered as the beginning and foundation of it. If that was not clear enough, Kant quite early maintained: “we have *two* duties of virtue. One duty of virtue to ourselves, and the second towards others” (P 27:163), anticipating the later position of the *Metaphysics of Morals* (cf. MS 6:487f). Not only

⁵⁶ *Ethica*, AA 27:737–32f. Therefore I disagree with Schneewind’s remark that the order of Baumgarten’s discussion of the ethical duties was “commonplace” (Schneewind 1997a, xxvii, n. 28).

⁵⁷ Cf., e.g., Crusius 1744, §§ 317ff, and Heineccius 1738, I, chap. V.

⁵⁸ On Kant’s relation with pietism see Kuehn 2001. Note that Siegmund Jacob Baumgarten’s main moral work, *Unterricht vom rechtmäßigen Verhalten eines Christen oder Theologische Moral* (1738), was composed in Halle while Alexander Gottlieb Baumgarten lived there, working on his *Ethica*: see Meier 1763, 12f; Gawlick and Kreimendahl 2011, xxiv.

⁵⁹ See, e.g., P 27:168ff after 135f; Kaehler 115ff after 61ff.

⁶⁰ Only the Vigilantius notes, anticipating the position of the *Metaphysics of Morals*, give up Baumgarten’s order and discuss the (alleged) duties to God after the ethical duties.

Baumgarten's priority of duties to God, but even the widely accepted idea that there are such duties, are thereby rejected.

Baumgarten's understanding of the duties to God relates also to a second major point of disagreement between Kant and him, since it entails, in Kant's eyes, an inadequate understanding of the duties to oneself. In general, Kant's remarks on the unfortunate lack of clarity on the grounds of self-regarding obligations (cf. P 27:187f; Kaehler 169f; C 27:340), directed against the general state of the discussion, clearly apply also to Baumgarten. Unlike Wolff, Baumgarten does not ground the duties to oneself on the goal of personal happiness (cf. Kaehler 171, 175; C 27:340f, 343), but his approach still focuses simply on the goal of perfection. The duties to oneself refer to various aspects of the perfecting of our faculties and aim at "strengthening the reality" of our soul, our body and our external state.⁶¹ Baumgarten seems to aim more at the general improvement of reality than of the agent as such (some formulations almost sound like anticipating Fichte's rephrasing the duties to oneself as duties *toward* oneself, and not for the sake of oneself) (cf. Fichte 1798, 246). Thus, if Baumgarten does not take the eudaimonist path, he still does not acknowledge what Kant takes to be crucial, that is, the status of moral subjects (cf. Kaehler 171f, 175; C 27:340f, 343; V 27:603).⁶² The apparently innocent remark that self-mastery "is the condition under which we can comply with all duties" (P 27:201; cf. Kaehler 203; C 27:360) rejects in fact Baumgarten's entire account, that confines the issue to a rather marginal observation (cf. *Ethica*, § 200).⁶³ As the comparison between the Powalski and the Kaehler-Collins notes shows, starting from this very point Kant develops the core of his new understanding of the duties to oneself as demanding "esteem for one's person" (Kaehler 203; C 27:360; cf. C 27:347).

Among the various aspects of Baumgarten's account that Kant rejects in virtue of his turn to a different understanding of the duties to oneself, I shall mention two examples. First, Kant is bound to reject the traditional subdivision in duties concerning different faculties of the soul and the body. Whereas in the Powalski notes Kant still seems to accept this organization (cf. P 27:202), the later lectures are unanimous in arguing against it, because the distinction between body and soul misses precisely "that which must be

⁶¹ *Ethica*, § 150: "Officia erga te ipsum sunt, quorum ratio perfectionis determinans est in te ipso ponenda realitas, sive propius animae, sive corporis, sive status externi realitates augeat." ["Duties to yourself are those whose determining ground of perfection is a reality to be posited in yourself, increasingly the realities either, more directly, of the soul, or of the body, or of your external state."]

⁶² Cf., e.g., also R #6590, 19:98.

⁶³ See further, more explicit critical remarks against § 200 of the *Ethica* in V 27:625.

determined through duty” (V 27:607, 625). Second, Kant repeatedly revises Baumgarten’s subdivisions, also arguing that some obligations that the tradition understands as other-regarding are in fact self-regarding, since the compliance or noncompliance with them affects the required “esteem for one’s person” more directly than the state of another person. Two prominent examples for this are lying and flattery.⁶⁴ Here again, thus, if Baumgarten’s noneudaimonist revision of perfectionism is certainly helpful, his account of the duties to oneself does not really provide a lead for Kant’s view, neither in the overall systematic, nor in the specific normative contents.

5 Final remarks

An investigation of Baumgarten’s significance for Kant’s moral philosophy cannot show any deep continuity between their views. Baumgarten does highlight some themes that become crucial in Kant’s ethical thought, especially obligation and the related concept of necessitation, along with the necessity of a noneudaimonist understanding of perfectionism. Still, their agreement on fundamental issues in moral philosophy is confined to very specific points. Furthermore, every one of these shared points is combined with serious objections of Kant’s against Baumgarten.

Indeed, Kant’s treatment of “his author” is mostly quite critical, as I have shown in a few significant examples. The carefulness and the philosophical significance of Kant’s criticisms of Baumgarten, however, should show how seriously he takes him. It should be clear that Kant’s comments on his statements are a substantial part of the development of his practical philosophy. Decades before Humboldt’s program, Kant’s lectures are in fact a prominent case of unity of research and teaching. He is clearly convinced of the philosophical worth of “his author,” and takes his views as seriously as those of Wolff and Crusius, Hutcheson and Rousseau. The critical dialogue that Kant entertains with Baumgarten is not of a different kind than the ones with these other authors, only closer and more thorough, if anything. Baumgarten’s role in the making of Kant’s critical ethics is clearly more pervasive than theirs, since he provides Kant with a general outline and countless terminological details for the elaboration of a comprehensive moral philosophy. In the first phase of its development, Kant’s ethical thought can even be described as the attempt to combine the structure of universal practical philosophy with the crucial innovations of Hutcheson

⁶⁴ On this, see Bacin 2013.

and other authors.⁶⁵ On Kant's view, Baumgarten does not provide solutions, but a clear outline of the subject and a valuable preliminary analysis of the basic concepts, helping to see what are the main issues to be discussed.

If Baumgarten does provide rich conceptual analyses, though, this is not always precise enough, in Kant's eyes, nor is it ultimately consistent or pointing in the right direction. In Kant's view, a consideration of the common use of reason in moral matters would have prevented Baumgarten from ultimately giving up a full objectivity of morality, or from forcing the ethical duties into a ranking and into distinctions that do not reflect the basic aspects of moral determination. Baumgarten proves indeed to be a one-eyed "technician of reason" in moral philosophy as well as in metaphysics, according to Kant's judgment that I mentioned earlier. Their views amount to two fundamentally different conceptions of morality, where the foundation of moral demands, God's role and the moral status of human beings are among the most significant points of disagreement.

⁶⁵ Cf. PS 2:298ff. For a reconstruction of Kant's project in ethics in the 1760s, see Bacin [2006](#), 1–66.

Herder: religion and moral motivation

Patrick R. Frierson

1 Kant and Herder

Königsberg in the early 1760s brought together, at formative moments in their careers, two figures who would go on to be among the most important philosophers of the late eighteenth and early nineteenth centuries: Johann Gottfried von Herder and Immanuel Kant. At this time, Kant was a “Magister” in his late thirties, in the midst of figuring out what kind of an intellectual he wanted to be. As a Magister, Kant could teach courses at the University of Königsberg but had to advertise his courses at his expense and received no salary from the university. His income was based on his classes’ popularity as he charged students fees. While he sought a professorship to provide a fixed salary and time for his philosophical work, Kant worked on developing his own philosophical stance and teaching interesting classes well.

Enter Herder, a seventeen-year-old, shy, rural student with amazing philosophical acumen and the style and diligence to excel.¹ From a poor family, Herder was discovered by a Russian army doctor who helped him get to Königsberg, where he received a small stipend teaching at Kant’s former primary school while attending the university (first for medicine, later for philosophy). Kant took an interest in Herder from the start, and Herder attended every course that Kant offered during the years he was at Königsberg. Kant allowed him to attend his courses for free (no small consideration for either of them) and came to identify Herder as his star pupil. Among the fruits of Herder’s attendance in those lectures is a set of lecture notes in ethics, metaphysics, logic, and even physics and mathematics.²

¹ This account of Herder draws from Zammito 2002, 138ff.

² These lectures notes are often incomplete. For portions of them, see AA 27:3–89 (ethics); AA 28:5–166, 843–941 (metaphysics); AA 24:3–6, 1099–1102 (logic); and AA 29:49–71 (math and physics).

All accounts of Kant as a lecturer during these years are glowing. Herder himself gives some of the best descriptions of what attracted him, and countless other students, to Kant's classrooms:

From his eloquent mouth flowed pleasant speech, rich in thought. A joke, a witty remark, humor, he used them always to good effect and at the right moment, remaining serious amid the general hilarity . . . His lectures were fascinating discourses, he would speak of an author and one could see his mind at work, probing further and further, and yet, in the three years I attended his lectures daily I never noticed in him the smallest sign of arrogance . . . I heard his evaluations of Leibniz, Newton, Wolff, Crusius, Baumgarten, Helvétius, Hume, Rousseau, some of them then very new names; it is remarkable that when he dealt with them his sole aim was a noble search for truth, a beautiful enthusiasm for everything that was best in man, a ceaseless, dispassionate desire to imitate what was best and greatest. (See Zammitto 2002, 142f)

At the same time, Kant's writings covered an exceptionally wide variety of topics and genres, from a satire on metaphysics to an elegant and popular essay on the feelings of the beautiful and sublime to a detailed "proof of the existence of God" to complex works in logic and natural sciences.

Strikingly, the end of Herder's time at the University of Königsberg coincided with the first offer of a Professorship to Kant. But this was not the professorship in logic and metaphysics Kant would accept six years later; it was a professorship in rhetoric and poetry, which he turned down. That Kant was offered such a professorship shows a significant part of why Herder was attracted to him. That Kant turned it down signaled the direction in which his philosophical development would progress, a development from more popularly oriented, observational, informal philosophizing, rich in feeling, to the rigorous, systematic, rational philosophy for which he became famous, a development that would take him further and further from his former student.

Elsewhere, I have discussed more general issues with Kant's moral philosophy during this period (Frierson 2012), and others have discussed the evolving relationship between Kant and Herder (Zammitto 2002). In the rest of this chapter, I focus on one particular topic: the role of religion in Kant's early ethics, particularly as that ethics is reflected in the lectures Herder transcribed,³ and particularly with respect to the problems of moral *motivation* with which Kant was particularly engaged.

³ For Herder's reliability as transcriber of Kant, see Zammitto 2002, 148f.

2 The problem of motivation in Kant's early ethics

During the 1760s, two of Kant's central concerns were ethics and the philosophy of religion. He lectured on ethics throughout this period, wrote on ethics in both a "Prize Essay" entitled *An Inquiry into the Distinctness of the Principles of Natural Theology and Morality* and his *Observations on the Feeling of the Beautiful and the Sublime*, and voraciously read whatever he could of the ethical writings of Wolff, Baumgarten, Hutcheson, Shaftesbury, Rousseau, and others. During the same period, Kant published his *Only Possible Argument for the Existence of God* (1762), *Dreams of a Spirit-Seer Elucidated by Dreams of Metaphysics* (a parody of rationalist metaphysics by means of comparing it to Swedenborg's spiritism, published in 1766), and an "Essay on the Maladies of the Head" (1764), which was, at least in part, a response to the increasingly popularity of a Polish religious fanatic who had recently appeared outside Königsberg. Both aspects of the philosophy of this early Kant reflected, as Herder put it, a "mind at work, probing further and further," rather than the (relatively) settled views of the Kant who would become famous for his "Critical" philosophy.

With regard to ethics, the views of the early Kant were preliminary in several respects. First, the Kant of the 1760s had not yet figured out the proper source for moral laws. As he puts it in his Prize Essay, "The ultimate fundamental concepts of obligation need to be determined more reliably" (PS 2:300). In seeking to determine the concepts of obligation, Kant was influenced by both rationalist-perfectionist moral theories (including Baumgarten's) that find moral content in rational principles, particularly the requirement to promote perfection (in oneself or others or both), and moral sentimentalists such as Shaftesbury and Hutcheson, who focus on moral feeling as the source of moral norms (PS 2:300). But religion – and in particular God's commands – might seem to provide another possible route to moral justification and content. Kant discusses this possibility during the 1760s as a philosopher still searching for a source of morality, rather than – as later – one who has settled on such a source.

Along with this concern about the nature and justification of moral laws, Kant is deeply concerned with the problem of moral *motivation*. Kant has not yet settled on what an appropriate and sufficient motivation for moral actions would be. He is drawn toward "moral feeling" as both a determinant of content and a moral motive, but he struggles with how to articulate the nature of this feeling. For instance, although he considers the possibility that natural benevolence or complaisance might be "raised to its proper

universality" (BGSE 2:216) and in that sense provide some motive, he is skeptical about the force of any such universalized sentiments, pointing out that such properly universal feelings are "colder" (BGSE 2:216): "He who loves the Tartar, loves not his neighbor" (H 27:67). Alternatively, Kant considers the importance of sympathy (see especially H 27:3) and even a repulsion from the "immediate ugliness" of vice (PS 2:300, see too BBGSE 20:24, 85, 93) as possible ways of making sense of the moral motive. Nonetheless, Kant is not wholly satisfied, and he continues to seek a proper formulation of how we can be morally motivated. In the absence of a well-worked-out account of "respect for the moral law" as the sole properly moral motive (e.g., G 4:400–402; KpV 5:71ff), Kant is open – at least in principle – to the possibility of religious motives for moral actions.

In addition to these general problems of moral justification and motivation, Kant is occupied in the 1760s with developing a properly *human* ethical theory, one that teaches "how properly to fulfill one's station in creation, and to rightly understand what one must be in order to be a human being" (BBGSE 20:41). In contrast to his insistence in the *Groundwork* that "a law, if it is to hold morally . . . must . . . not hold only for human beings" (G 4:389), Kant emphasizes in the Herder lectures:

the *highest morality* is not on a par with the *moral level* of man . . . [W]e should investigate the degree of morality that is suited to human beings . . . An ethic for the human being, *determined* in his nature, by his knowledge, powers and capacities, has yet to be written. For by reason we can also discern rational perfections that are suitable, indeed, for a higher being, but not for him. (H 27:61f)

This problem is relevant to both moral content and moral motivation. With respect to the latter, for example, the problem with universalizing benevolence is not a problem *in principle*, but rather for beings *like us*. Thus, Kant aims for a *realistic* account of moral motivation, which leads to a concern with humans' moral imperfections, particularly that we are assailed by temptations to which we frequently succumb. Kant seeks an account of the moral motivation that can strengthen us against those temptations.

Moreover, with respect to motivational issues, Kant further narrows his interest by focusing not only on human nature in pristine purity but also on human beings *in society*. In this respect, Kant is heavily influenced by Rousseau, even while explicitly contrasting his approach with Rousseau's: "*Rousseau* . . . starts from the natural human being, I . . . start from the civilized one" (BBGSE 20:14). And this starting point leads to even more pressing problems with moral motivation because the temptations for the

civilized human are, as Kant emphasizes throughout the Herder lectures, more pressing than for the natural human: "In the savage state, instincts are enough . . . In civil society, where the needy have multiplied, . . . feeling is much weakened . . . [and t]he civilized man is much constrained by self-serving artificial desires" (H 27:58).⁴ One result of the focus on the civil condition is an increased role for religion. For Kant, religion is not necessary for morality: "the Hottentots . . . possess moral feeling" (H 27:11) and "in the state of nature there is less occasion for religion . . . [In] the first state of things . . . [people] are . . . morally good without God" (H 27:76). Kant condemns the moralist who "presupposes comforts, honour, etc., though that is unnatural; [and] extends duties contrary to nature" (H 27:63). Such people "fabricate false virtues" because "those that are *appropriate* to the *natural* man are too *elevated* and hyperbolic for the *artificial* one" (H 27:64). In some cases, the motivational problems of the civil condition can arise *from* religion, as when one "derives marriage not from the sexual impulse but from the command of God" (H 27:63f). But Kant sees importance for religion in the context of the civil condition's increased moral temptations:

when many of his comforts turn into necessities, [the human being's] impulses gain the upper hand, so that morality becomes too weak, and the religion of nature does not suffice. For this, more understanding and philosophical reflection are required, than can be expected of the whole human race. So it has to be complemented by a revelation. (H 27:74f)

Religion, even in the form of special revelation, is needed to counteract the morally weakening effects of the civil condition.

For the early Kant, then, moral motivation is a serious problem. Most fundamentally, Kant is trying to figure out the nature of moral motivation in general. Having not yet isolated "respect for the moral law" as the sole moral incentive (see G 4:400), Kant needs an account of what moral motivation *in general* is like. But Kant is also occupied during this period with thinking through a properly *human* ethics. And, under the influence of Rousseau, he is particularly concerned with motivational problems arising from the luxury, excess, and moral corruption of "civilization." For all of these motivational issues, Kant turns to a variety of different possible incentives to moral action, including certain feelings such as sympathy or

⁴ In this context, Kant suggests the need for rational bases of morals: "pity is here replaced by the concept of what is *right* . . .; here virtue becomes calm and rational, and no longer remains a mere animal instinct" (H 27:58). But the same problems in the civil condition that lead Kant to shift from pity to reason also lead him to an increased role for religion in moral motivation.

complaisance, an innate aversion to moral ugliness, and – of central concern in this chapter – to various forms of religious motivation.

In the next two sections, I focus on several positive roles for religion. First, although religion will not be *the* moral motive, it can provide a “new motivating ground . . . derived from God’s *arbitrium*” (H 27:11), which can *directly* cooperate with more purely moral motives. Second, religious motivation can be a “preparatory” motive conducive to ethics (H 27:14), particularly in the context of the motivational problems caused by the moral limits and failings of civilized human beings. This preparatory function provides the framework for understanding two indirect roles of religion in moral motivation. Religion helps cultivate the resignation required to dampen the corrupting influence of luxury on desire, and if considered properly, it prevents moral despair without fostering lazy self-confidence. In the [next section](#), I take up the direct role of religion in moral motivation. In [Section 4](#), I discuss the two indirect ways it prepares for morals. And in my conclusion, I briefly discuss some of the ways that religion goes wrong.

3 Religion and moral motivation 1: Direct religious motivation

As we have seen, Kant in the 1760s has not yet settled on a definitive account of moral motivation, so there is at least some room for considering religion as *the* moral motive. However, even before he has settled on “respect,” Kant makes clear that religion cannot provide the legitimate *primary* motive for fulfilling moral obligations: “the ethical motivating grounds should always be moral . . . [and] must be drawn solely from noble, virtuous and free choice” (H 27:14). Thus, religion has nothing to offer in terms of satisfying Kant’s desire for a clear explanation of the nature of moral feeling. In this sense (and, as we’ll see, in another as well), religion is strikingly similar to two “sort[s] of kindly feeling which [are] to be sure beautiful and lovable but still not the foundation of a genuine virtue” (BGSE 2:216). As discussed in his *Observations* (published at the same time as he was giving the Herder lectures), these feelings – sympathy and complaisance in particular – can be “weak” and “blind” and, when present as one’s primary motives, “only contingently agree with” virtue (BGSE 2:215). Thus, although beautiful, neither sympathy nor complaisance – nor, it turns out, religion – can solve Kant’s problem of discerning the primary moral motive.

To get clearer on why religion cannot function as the primary moral motive, it is necessary to get a sense for Kant’s attitudes during this period toward the relationship between moral obligations and divine commands. Importantly, at this stage, Kant could in principle be open to a divine

command theory of moral duties. In these lectures, Kant asks explicitly, “Can we, even without presupposing God’s existence and His *arbitrium*, derive all obligations from within?” (H 27:9) and “Should Christian ethics precede the philosophical or vice versa?” (H 27:15⁵). But the Kant of the Herder lectures – like the Kant of the *Groundwork* – already answers the former question “in the affirmative” (H 27:9), saying that “the *natural* must justly precede” (H 27:15). Kant’s justification for the priority of nonreligious grounds for obligations hinges on his distinction between what is good “in a physical sense” and what is genuinely morally good. Physical goods are conducive to our physical well-being but not in themselves morally good. Given this distinction, Kant explains several related but distinct reasons that the content of our obligations must be derived “from within”⁶ rather than from God’s choice. These reasons depend on the fact that “From the *arbitrium divinum* I cannot myself obtain the relevant concepts of the good, unless the concept of the morally good be assumed beforehand” (H 27:9). Thus even “the judgment as to the perfection of God’s *arbitrium* presupposes the investigation of moral perfection. Otherwise, the sheer *arbitrium* of God is good merely in a physical sense” (H 27:9). That is, without a prior standard of goodness, God has only the “good” of a strong ruler or a Creator, neither of which is intrinsically moral.

Kant goes further, however, highlighting the moral impotence of purely religious motives:

Supposing the *arbitrium* of God to be known to me, where is the necessity that I should do it, if I have not already derived the obligation from the nature of the case? God wills it – why should I? He will punish me; in that case it is injurious, but not in itself wicked; that is how we obey a despot; in that case the act is no sin, in the strict sense, but politically imprudent . . . Without the prior assumption of obligation, punishments come to nothing; what God displays is merely ill-will. (H 27:9f)

Even if we knew that God’s will was morally perfect, we would need some concept of obligation that tells us to obey morally perfect wills, lest our obedience, even to a morally perfect God, be merely obedience as to a despot, obedience out of physical but not moral grounds. Moreover, even

⁵ Translations of portions of the Herder notes not included in the Cambridge edition are based on a draft translation of those notes undertaken by Patrick Frierson, Karah Kemmerly, and Luke Rodriguez.

⁶ The “within” that these obligations are derived from is not yet pure practical reason. In these early lectures, Kant discusses knowledge of obligations coming from “the nature of things” (H 27:7, 8), “moral feeling” (e.g., 27:4–6, 10, cf. 2:300), and a direct awareness of moral perfection (H 27:13). For further discussion, see Shell and Velkley 2012.

the notion of “obligation” itself depends on human beings having a volitional structure susceptible to intrinsically normative necessitation. As Kant explains:

The revealed ethic, if it is to be practical, must ground itself upon the motives of the *natural* ethic. Like any revelation, it presupposes natural powers, e.g. capacities of the soul that are fit for the purpose. Otherwise, it would be at most a miraculously transforming book; but in fact it is a book that *lays obligations* upon us, and presupposes instruments and receptivity in the face of revealed religion. (H 27:16)

Here Kant is considering a divine influence on behavior that need not be mediated by pursuits of physical goods but might even be the result of “divine inspiration and influence” (H 27:10) such as through a holy book. But in whatever way God were to influence a person, unless such an influence proceeded by means of *natural powers* conducive to a sense of obligation, it would not *obligate* but only control. For Kant, then, obedience to God can be morally good only when it is *secondary* to a more fundamental “deriv[ation] of all obligations from within” (H 27:9).

Nonetheless, despite all his arguments that internal grounds of obligation precede divine commands, Kant argues that religion in general and a sense of divine commands in particular are morally important: “from [our concept of obligation] we come to God’s power of choice” (H 27:9). Although we must formulate a concept of obligation based on internal grounds alone, we can and should go on to see our obligations as originating from God’s will. This conception of moral laws as also arising from divine commands is consistent with Kant’s claims in his *Critiques* (see KrV A 819/B 847; KpV 5:129; KU 5:481) and with his claim in *Religion* that although “morality in no way needs religion,” still “morality leads inevitably to religion” (RGV 6:3, 6). In the Herder lectures, however, this emphasis on “natural religion” is more pronounced, and Kant details why the “application of the divine arbitrium . . . as a ground” of moral obligations constitutes “a part, but not the basic principle, of morality” (H 27:9).⁷

One possible argument for extending natural morals into natural religion is based on God’s omnipotence as Creator: “since God . . . is the ground of all things, this is also the case here” (H 27:9). Just as God creates the physical world, God is the source of moral obligations. But Kant limits the importance of the argument from omnipotence by siding with those who see

⁷ In these lectures, Kant even entertains the idea that religion might give rise to duties (such as a strict prohibition on suicide) derived from “a *special* divine will” (H 27:30, see, too, H 27:84).

God's *nature* rather than God's power of choice as the ground of morals: God "is indeed the ground of it, but not *per arbitrium*" (H 27:9). Rather, "since He is the ground of possibility, He is also the material ground (since in Him all things are given) of geometrical truths and morality. In Him there is already morality, therefore, and so His choice is not the ground" (H 27:10, cf. BGSE 2:100f). Thus, the argument from God's creative omnipotence is *not* Kant's primary argument for recognizing the importance of divine *law*.

Consistent with his later rejection of rational-theological proofs of God in favor of morally practical proofs,⁸ Kant replaces what is a primarily metaphysical-theological basis for grounding ethics in divine commands, with a moral-practical one:

morality is incomplete, if all grounds of obligation are not included, . . . [s]o the *arbitrium divinum* should never be left out . . . [since then] our moral perfection becomes incomplete, if it arises solely from inner morality, and is considered without reference to God's *arbitrium*. In the absence of the latter, my action is already still moral, indeed, but not so completely good, morally, as when it conforms to all grounds. Those who attend solely to the *arbitrium Dei* are considering merely their liability to the *jus naturae divinum*; but we should attend also to the inner morality, and consider obligation as well. *Ethica rationalis*: the one without the other is not universal morality, and indeed far less than this; we are virtuous already from the nature of the case, pious only in having regard to the *arbitrium divinum*. (H 27:10)

At this stage in his reflections on moral philosophy, Kant still allows that moral perfection can have multiple, irreducible elements. For better understanding how this notion of "all grounds of obligation" works in the context of religion, one key difference between Kant's early ethics and his later is particularly salient. Although Kant's moral philosophy from the *Groundwork* on emphasized the importance of *purity* in moral motivation, no such emphasis was present in his early ethics. Thus the possibility of cooperating moral motivating grounds was discussed throughout these early works. In the *Observations*, for example, even while rejecting sympathy and complaisance as *the* moral motive, Kant goes on to insist that these feelings are genuinely "moral [*moralischen*] qualities" (BGSE 2:215) that have a "kinship" with virtue (BGSE 2:217). Moreover, "to the extent that they harmonize with virtue, [they] may also be regarded as noble" (BGSE 2:215) and when they are properly proportioned, they even "bring

⁸ But compare his *Only Possible Demonstration of the Existence of God*, published in 1762–1763 (AA 2:63–163).

about the noble attitude that is the beauty of virtue,” warranting for themselves the title “adopted virtues” (BGSE 2:217). In sharp contrast with the emphasis in the *Groundwork* on displaying duty in all its purity, Kant is interested throughout his early ethics in showing how virtue as a whole, although it requires a purely moral motivating ground, is enriched by quasi-moral qualities that are part of completed human virtue.

In the Herder lectures, this importance of combining the pure moral motive (whatever that will turn out to be) with other motivating grounds is particularly important with respect to religious motivations. Just as one who is virtuous but also sympathetic has a beautiful virtue, in the religious case, Kant allows that although someone who does their duty for its own sake is “already still moral,” someone who *also* does it because God commands it is better, and better in a way that can be called “moral.” Both rational and theological ethics are incomplete when isolated:

To disregard the one [the inner grounds of obligation] is wicked; to disregard the other [divine commands], godless; the former are moral errors, the latter, sins; the former concern the moral teacher, the latter, the preacher; the one wishes to have people morally good, the other wants their moral goodness to be complete. (H 27:10)

Just as sympathy and complaisance make virtue kindly, beautiful, and lovable, attention to moral laws as divine commands makes virtue pious and reverential: “Thus in their highest stages, all morally good actions are religious acts” (H 27:17).

The only indication in this particular part of the lectures as to how, precisely, recognizing moral obligations as divine commands “completes” ethics is a repeated emphasis on the divine will as an *external* ground of duty: “the *arbitrium divinum* is a ground of external obligation for our morality, so the *arbitrium divinum* should never be left out, as an external obligating ground” (H 27:10). But given his prior argument against obeying God merely for “physical” reasons, it is not clear why this external ground should be necessary or even helpful for completing one’s moral perfection. Kant even warns that “He who has a notion of the external obligation, without the inner, sees the motivating grounds as tasks, which do not make him moral at all, but merely politically crafty” (H 27:10).

One explanation for the importance of adding religion to morals that is more consistent with Kant’s overall account derives from the importance of morality as the proper means of relating to an all-good God and more generally toward harmony with morality in the world as a whole. Kant explains:

this new and higher morality . . . contains a relationship [*Verhältnis*] to the greatest supreme rule, which is the ground of everything, and thus constitutes the greatest harmony . . . I must first abstract my actions from the divine will, in order even to recognize the goodness of that will. But once I have perceived it with sufficient abundance, exactitude and vividness, it becomes the supreme basis, (1) because the knowledge is then noble, and (2) because it provides the highest degree of vividness . . . When a man then rises to the highest level, that shows him as God's supreme instrument. (H 27:17f)

Kant is thoroughly suspicious of grounding moral obligation on divine commands in a way that would make obedience to such commands a matter of fear of punishment. Hence, even this passage is *primarily* focused on arguing that "this new and higher morality is only brought in afterwards" and thus that "our author's [Baumgarten's] method is incorrect, since it begins from religion, whereas it ought to have started from a morality, which would then be increasingly purified" (H 27:17f). But Kant also suggests a different sort of obedience to the divine will, one aiming for a harmonious cooperation with God. One first recognizes what is morally good, then turns to a recognition of those obligations as flowing from the good will of a wholly good God. And then one obeys that God not out of fear of physical punishment but out of awe and love toward a supreme goodness whereby one comes to see oneself as an instrument of that supreme goodness.

In this context of obedience to God as a moral motive, Kant works out, in some detail, different sorts of "fear of God," explaining that one can (and should) "be moved by a morally perfect sublimity" while pointing out that this "reverence . . . can actually suppress love" in some contexts. Some have a "*fearfulness towards God [Furcht vor Gott]*" that is "*servile*" and "recoils from actions because of punishment" (H 27:32f). This servile fear "diminishes love; for as soon as we see somebody against us, a degree of love is eliminated; who loves anyone, insofar as he punishes us?" But one can, instead of this fearfulness, have a proper "*fear of God [Furcht Gottes]*" (i.e., awe) . . . that can coexist with love, . . . [and] is thus an *awe coupled with love*." In that context, "We guard against God's displeasure because of His beautiful and sublime qualities; but not, to that extent, from fear; it is ourselves we are in fear of, for we would . . . be hateful in our own eyes" (H 27:31f). One who properly understands the goodness of God seeks to avoid actions that displease God, not from fear of possible punishment by God, but out of that awe mixed with love that would find it hateful to displease God. The result is a will that does what is good *because* it is good

but *also* because it is the good decree of a good God with whom my will is in “harmony.” Insofar as we raise moral feeling through religion, we not only reach “the highest degree of such sensation” of God, but we also attain “the whole *summa* of morality” (H 27:19).⁹

Raising morals to the level of religion, then, adds piety to morality, putting one who acts out of moral motives as well as obedience to divine commands into harmony with supreme goodness and a right relationship of awe and love for God. Although one could, in principle, act merely from inner grounds of moral motivation without any regard for God’s will, such actions would be “godless” or “sins” and thus one’s moral goodness would be incomplete (H 27:10). Incorporating religion into one’s motives recognizes moral laws as not *merely* natural laws but also positive laws of a supreme lawgiver. Kant exhorts, “Try especially to always couple the idea of God with your morality; first with your *natural* moral feeling, so that your immediate liking for the good becomes, in the light of God, religion. Try also to make the idea of God *dominant* in the *depths* of the soul” (H 27:23).

4 Religion and motivation 2: Indirect religious motivation

The last section emphasized the direct quasi-moral importance of being motivated by religious considerations, but for Kant, religious motives are particularly relevant given human moral limitations,¹⁰ which lead Kant to several indirect ways religion functions to combat these imperfections. In discussing “external” motives, Kant explains that although “ethical motivating grounds should always be moral and not merely practical as physical means,” nonetheless these merely “subjectively motivating grounds” can “become *mediately* motivating grounds . . . [and] are very good, and often *preparatory to ethics*” (H 27:14, second emphasis added). Again, the case with

⁹ In line with his emphasis on moral feeling during this period, Kant even argues that moral feeling – which he identifies with “conscience” – is all that is left of what was once “the immediately clear sensation of the divine presence,” a sensation that was “not symbolic, but intuitive; not from inference but from sensation.” As he explains:

With us, perhaps, the broadest and vaguest concept thereof still resides, even now, in conscience. If we directly improve our moral feeling, we approach the divine presence in sensation; so maybe such people again develop the image [of God], although their spiritual utterances sound fanatical; and religion elevates us to the highest degree of such sensation . . . This is true morality; a part of it already precedes all religion, but a part is greatly enhanced by religion. (H 27:19)

¹⁰ See, e.g., H 27:9: “For one who has not wholly fulfilled his obligation, morality is incomplete, if all grounds of obligation are not included” (emphasis added).

religion in the Herder lectures parallels Kant's earlier discussion of sympathy and complaisance in his *Observations*, where he noted:

In recognition of the weakness of human nature and the little power that the universal moral feeling exercises over most hearts, providence has placed such helpful drives in us as supplements for virtue, which move some to beautiful actions even without principles while at the same time being able to give others, who are ruled by these principles, a greater impetus and a stronger impulse thereto. (BGSE 2:217)

In the Herder lectures, Kant even more clearly connects the need for preparatory motives with human imperfection:

Everyone, to be sure, has need, in part, of *sensitive jucunda* [sensitively pleasing], and in part of *sensitive molesta* [sensitively displeasing], even for *moral* actions. For our moral feelings are so buried away under the sensuous, and the sensory motivating grounds thus make it easier for the soul subsequently to make its decisions on principle. By those principles which outweigh the sensory motives, we are brought nearer, as it were, to the domain of morality. This extends, not merely to the teaching of ethics, but also to education and religion. (H 27:14f)

Like Kant's much later discussion of respect for the moral law (in the *Critique of Practical Reason*), moral motivation involves a suppression of sensory motives for the sake of morally good action. But whereas the respect of the *Critique of Practical Reason* partly *consists in* the humbling of sensory motives, in these earlier lectures, this suppression is *preparatory* to moral motivation and can be the result of other sensory motivating grounds. Thus, Kant considers various motives, including religious ones, as not-strictly-moral means by which one outweighs sensuous motives in order to give properly moral motives greater influence. In addition to functioning as a direct motive to morally good actions, then, religion can serve as an indirect preparation for morality that "makes it easier" for properly moral motivations to function. Although Kant discusses several indirect roles for religion, this section focuses on only two of the most important: the cultivation of resignation and the alleviation of moral despair.

One of the main problems with the civil condition is the proliferation of "self-serving artificial desires" (H 27:58) where "comforts turn into necessities" (H 27:74), so one of the most important roles of religion within Kant's early ethics is the cultivation of a resignation to God's providential care that loosens attachments to objects of particular desires by shifting focus toward the good of the whole. Kant explains, in detail, how the civil condition erodes morally proper love for fellow human beings through the

generation of “fanciful desires”: “the more there is of excess [*Ueppigkeit*], the less there is of practical human love. For by excess we multiply in fancy our own needs, and thus make practical love *difficult*” (H 27:64). In response, Kant comments that “a person does not actively *love* another *until he is himself in a state of well-being*; since he is the *principium* of the other’s good, let him first better himself” (H 27:64). Following the guidance of Rousseau (and earlier Epicureans), Kant rejects the model of “bettering himself” that would multiply *resources*. The attempt to put oneself in the condition to satisfy one’s luxurious desires is ultimately self-defeating. Instead, “To make itself practical, [one should] puts oneself *at ease* with oneself, *making do with little*; and from this comes practical love . . . hence in a condition of simplicity there will be much practical love, and in a state of luxury, little” (H 27:65). To prepare the way for genuinely moral motivation, one of the most essential steps for civilized human beings is to clear away the false sensuous incentives arising from fanciful excess.

For shifting attention away from fanciful desires and providing the “ease” needed to make love practical, religion can provide invaluable assistance. Kant claims that “resignation to the divine will is necessary” and really constitutes the proper form of “love of God” (H 27:26). In expanding on the nature of this resignation, he explains:

Insofar as we regard all acts of God as the best means, to the best end, of happiness, we may be completely at ease. The great mutability of things, and the storms of my passions, can best be comforted by the thought that I am placed in the world, and placed there by supreme goodness, not for my own benefit; and however uncertain the order of nature may be, it is nevertheless under the supreme being; and in this way, then, *religion* alone may be completely reassuring. For even a naturally good and moral man must always tremble before *blind fate*. (H 27:25f)

The importance of proper religion for resignation and calming the passions leads Kant to emphasize the nature of proper trust in God (H 27:28f). We are to trust in God’s goodness for “the whole” rather than call on God to provide particular objects of present desire:

True trust is thus always connected to self-denial, as I am always convinced [that] I could not determine the measure of his goodness in an individual case, and the *greatest trust* is even connected with the *greatest resignation* because I still always trust God for the greatest goodness in the whole, although I don’t venture to specify this goodness in a particular case. (H 27:28)

For human beings naturally tempted to subordinate moral laws to sensuous temptations, and particularly by those in the civil condition with exaggerated

and unnatural desires and a diffuse and weakened moral feeling, the shift of attention involved in genuine religion provides a healthy preparation for morality. Through genuine trust in God, one learns that the satisfaction of particular desires is not the highest good – “eternity serves . . . to diminish the ills of this world” (H 27:44) – and one can thus allow oneself to be motivated by moral feeling.

A second important motivational problem, one present in morally imperfect human beings in general but particularly exacerbated by a civil condition that weakens moral resolve while heightening moral demands, is the problem of moral despair. Kant puts this point particularly clearly in the context of an argument for why the natural ethic must precede the Christian. There he explains:

the natural ethic shows us many obligations which are impossible *secundum quid*, and thus leads to the Christian ethic; the former creates the contradiction in man, that he *imputes* to himself something which he cannot omit; it creates the collision between impotence and the moral ordinance, which the Christian ethic reconciles. (H 27:15f)

Kant’s immediate point, in context, is that we cannot make sense of why the elements of the Christian ethic (presumably various means of grace) are necessary without antecedently understanding humans’ incapacity to fulfill the obligations of the natural ethic.¹¹ The corollary to this claim about the logical priority of natural ethics, however, is a point about the necessity of Christian ethics to “reconcile” the “collision” between our capacity and morality’s demands. Kant does not lay out in detail how the Christian ethic solves this problem, but a later comment provides some clue: “a deficiency of confidence that God will improve our morality can make us [either] despairing or way too reliant on our own powers without dragging God into it” (H 27:29). Arguably, what Christianity¹² can contribute to alleviating moral despair is the clear sense that God can and will improve our morality. This suggestion is confirmed by Kant’s later claim about the proper focus of appeals for divine aid:

The highest degree of connection with God as a means is when we utilize the divine will as a means to the betterment of our own morality. Julie says, for

¹¹ In this sense, he appeals back to the traditional theological argument-structure of Paul’s Letter to the Romans, where Paul uses the fact that the Gentiles have “the law written in their hearts” to justify the claim that “all have sinned” and must be “justified . . . by Christ” (Romans 2:15, 3:23f).

¹² To what extent Kant would apply this to other religious systems is difficult to discern in the Herder lectures. Kant often emphasizes tolerance and religious pluralism but also seems to see some advantages of Christianity in particular.

example: Our good actions are noticed by witnesses: – she uses God’s will to better her morality; but to use it merely as a means to happiness is ignoble, and no religion. (H 27:18)

Just as in the case of providence and resignation, so here the role of religion is to provide aids to morally good actions. Julie¹³ does not, importantly, see God as a rewarder of happiness. Rather, she sees God as an aid to her own moral development, whether as an ever-present witness or, as Kant suggests later, as one who can “increase our moral perfection,” albeit not without our own effort (H 27:29). This awareness that God is present with assistance towards moral perfection can protect weak human beings from both arrogant complacency that ignores our own moral shortcomings and the despair that simply accepts them.

5 Conclusion: Moral dangers of religion

The Kant that shows up in the Herder lectures is deeply concerned with questions about religion and about moral motivation, and these come together into a theory of moral motivation that accords religious motives at least three important roles. Religion functions as a direct moral motive, one that must be combined with a purely moral internal motivating ground and that, when so combined, helps to complete morality by adding a further good motive, one that makes actions both good *and pious*. And religion functions indirectly as a preparation for pure moral motivation, both by cultivating resignation and by alleviating moral despair. Both of these indirect roles are particularly important for human beings corrupted by the luxury of the civil condition.

For Kant, of course, religion is not without its dangers, and much of Kant’s religious thought in the 1760s is focused on moral dangers arising from bad forms of religion. Kant points out, for instance, how Julie’s noble supplication for divine assistance can become a “lazy trust” that “trusts God, so far as to despise external means” (H 27:29) and thus undermines rather than cultivating moral virtue. Unlike the admirable pleading of Julie for God to cooperate with her own efforts toward moral perfection, lazy trust expects God to take care of everything for me, and in just the ways I want and expect (see, too, H 27:30). Likewise, “the *moralist* or *cleric*” can exacerbate rather than alleviate the problems of the civil condition when they “presuppose comforts, honour, etc., though that is unnatural” or “extend

¹³ See Jean-Jacques Rousseau’s *Julie, or the New Heloise*, published in 1761 and widely read (including by Kant) (Rousseau 1997 [1761]).

duties contrary to nature, e.g. by deriving marriage, not from the sexual impulse, but from the command of God” (H 27:63f). Here religion actually cooperates with the civil condition in both extending the range of our luxurious desires and in placing moral demands that are unnecessarily difficult to fulfill. And much of Kant’s discussion of the value of religious resignation focuses on exposing the dangers of “uses of religion not just for the chief end of religion [virtue] but also for a minor subordinate end” such as the satisfaction of particular desires (H 27:35).

Beyond these particular issues, the Herder lectures show significant preoccupation with four main ways in which religious thinking was going astray in his time. All of these reflect failures to recognize the proper role of religion as supplement to virtue, but in very different ways. Two – religious fanaticism/enthusiasm (e.g., H 27:21–4, 40) and religious intolerance (e.g., H 27:73–8) – reflect a failure to sufficiently attend to natural morals as the primary basis of obligation, *replacing* morality with God rather than *supplementing* it with true religion. A third – empty theological speculation (H 27:19f, 22) – focuses on theoretical claims about God without seeing the primarily *practical* import of religion. And a last – atheism – can reflect mere thoughtlessness about God, an excessive speculation, or a “moral ground,” which is “very dangerous to society” (H 27:11, see, too, H 27:37f, 80). All of these issues are discussed with care in these lectures, and it is during this period that Kant begins – in the *Remarks* – to develop a notion of metaphysics as “a science of the limits of human reason” that “do[es] not remove useful certainty, but useless certainty” and “removes the appearance that can be harmful” (BBGSE 20:181). This general usefulness of metaphysics eventually comes to the fore with respect to the very religious threats Kant worries about in his early work. As he puts it in the *Critique of Pure Reason* over fifteen years later, “Through criticism alone can we sever the very root of **materialism**, **fatalism**, **atheism**, of freethinking **unbelief**, of **enthusiasm** and **superstition**, and finally also of **idealism** and **skepticism**, which are more dangerous to the schools and can hardly be transmitted to the public” (KrV B xxxiv).

*Collins: Kant's proto-critical position**Manfred Kuehn***I Introduction**

The lecture notes “Collins,” published in the Academy edition of Kant’s works under the title *Moralphilosophie Collins*, are based on a notebook of one Georg Ludwig Collins (1763–1814), who seems to have attended Kant’s lectures during the winter semester of 1784–1785.¹ The title page of the notebook reads: “Moral Philosophy, in accordance with the lectures of Professor Kant, Königsberg, Winter Semester 1784–5, Georg Ludw Collins, d.GG:S. on Baumgarten” (C 27:241).² On the last line of the notebook, we find: “*Finis*, Königsberg, April 19th, 1785” (C 27:471). None of this means that the notes were actually taken by Collins himself during that time period. They just indicate the time frame during which Collins copied the notes from a preexisting copy of that notebook. Because copies of lecture notes of the *collegia* given by the full professors at the university could freely be bought in Königsberg, it is virtually certain that Collins bought a ready-made set of notes.³ The best estimate for when the lectures on which they are based were held is some time around 1775.

In any case, the Collins lecture notes belong to a group of thirteen different sets of lecture notes which are closely related to each other (Stark 2004b, 392, 402ff). Some of the most important of these have been published before. The best known and most widely available is the edition

¹ They are to be found in volume 27.1, 241–471, of the Academy edition, that is, in *Kants Vorlesungen*, volume 4. English translation in Heath and Schneewind 1997, 37–222. Collins was inscribed at the University of Königsberg on October 9, 1784, as a student of theology. He stayed for only one year. In 1785 he went to Leipzig, and then later to Riga. See Lehmann 1979, 1050.

² I have changed the “from” in the translation to “in accordance with” because “*nach*” cannot mean “from” here. Nor did Kant read *from* Baumgarten. He used him as a guide for his lectures. The abbreviation “d.GG:S” identifies Collins as a theology student (“*der Gottesgelehrtheit Student*”). Except where indicated otherwise, all translations are my own.

³ For the decisive argument, see Stark 2004b. Collins seems, however, to have copied the set of lecture notes himself, probably using this exercise to review the material, and he probably added marginal comments for similar reasons. See Lehmann 1979, 1060, for instance.

published by Paul Menzer in 1924 under the title *Eine Vorlesung Kants über Ethik* (see Infield 1930 for an English translation). This is based on three different sets of notes which are in all likelihood copies of one common *Urtext* (see Menzer 1991; cf. Schneewind 1997a, xv–xix). Menzer actually used three notebooks in preparing his edition. The copy text is based on Theodor Friedrich Brauer's notes, but he also used a set of notes by Gottlieb Kutzner and Ch. C. Mrongovius to “check and correct” the manuscript (see Schneewind 1997a, xvi). For Menzer, there was no doubt that these three notebooks were all copied from one *Urtext* and that the contents present “a substantially accurate transcription of Kant's lectures as he gave them during the years 1775–1780.” Furthermore, he surmised that they probably represent the form the lectures took during “the later part of this period.”⁴ Menzer appears to have consciously decided against using the Collins notes as one of the basic texts for his edition. Gerhard Lehmann, the editor of volume 27 of Kant's Academy edition, by contrast, appears to have selected the copy of Collins *just because* it was different from Menzer's three copies (of which two were not available at the time when volume 27 was prepared; see Lehmann 1979, 1041). It may be doubted that this was a good decision.

Another text which is very important for the Collins notes is a notebook by one Johann Friedrich Kaehler, who was a student of jurisprudence at Königsberg. Kaehler's text has been published by Werner Stark (2004a). It represents in many ways the best edition of the three lectures on moral philosophy available today. Also very closely related to the Collins notes, this set was probably based on lectures held either in the winter semester 1773–1774 or in the winter of 1774–1775. Kaehler's notes may very well represent the *Urtext* on which all the other relevant lecture notes, including those of Collins, are based. But, however that may be, they are certainly closer to the *Urtext* than any of the other sets of lecture notes.

Kant almost always used two textbooks by Alexander Gottlieb Baumgarten in his lectures, namely, the *Initia philosophiae practicae primae* of 1760 and *Ethica philosophica* (1740, 1751, 1763).⁵ Like Baumgarten, Kant made a sharp distinction between “universal practical philosophy” and “ethics” proper. The first part dealt with the principles of practical philosophy in general, whereas the second part investigated the broad outlines of ethics. This distinction is reflected by the fundamental division of the

⁴ Schneewind 1997a, x and xvi. Menzer said “1775–1780,” not “1775–1784” as related in Schneewind 1997a, xvi.

⁵ They are reprinted in the Academy edition, AA 19:4–91 (*Initia*); AA 27:733–1015 (*Ethica*).

course. Accordingly, the text of the Collins lectures is divided into two parts of unequal length, the first and shorter one dealing with the general principles of practical philosophy and ethics, the second treating substantive questions of ethics.⁶ It would not be entirely inappropriate to say that the first part corresponds roughly to what Kant later tackles in the *Groundwork of the Metaphysics of Morals* and the *Critique of Practical Reason*, whereas the second part deals with the subject matters relegated to the *Metaphysics of Morals*. This correspondence can, of course, be approximate only because Kant, in 1775, had not clearly developed his own conception of ethics. Still, it is surprising how much light the lectures throw on later developments and how far the correspondences go. First of all, the conception of a “highest principle of morality” is central both in the Collins part on universal practical philosophy and in the *Groundwork*. Kant claims that the *Groundwork* is actually “nothing more than the search and establishment of the *supreme principle of morality*” (G 4:392) and he claims that this is a task that is complete in itself and can be kept separate from “every other moral investigation” (G 4:392). A similar claim may be made about the part on universal practical philosophy. Although he does discuss some of the parts of Baumgarten's *Initia*, he is, as we shall see, most centrally concerned with “the highest principle of morality” (C 27:274).

Accordingly, I will divide the remainder of my discussion of the Collins notes into four parts. The first (Section 2) will deal with the relation of the section on ethics or “*Ethica*” in the Collins lectures to the *Doctrine of the Elements of Virtue* of the *Metaphysics of Morals*. The discussion of the “Ethics” will be relatively short, trying to establish some of the general features of Kant's ethical outlook in order to determine how far his ethical outlook changed over the years. The second part of my discussion (Section 3) will primarily discuss Kant's account of “The Ethical Systems of the Ancients” and what this says about the relations between religion and moral philosophy, as well as those between empirical and rational accounts of morality. The third part (Section 4) will concentrate on “Universal Practical Philosophy” and its relation to Kant's mature doctrine of the categorical

⁶ It is inexplicable to me why the editors of the Cambridge edition of Kant's *Lectures on Ethics* divide the lecture into three main parts, namely, on “Universal Practical Philosophy,” “[On Religion],” and “On Morality.” In the German text, the chapter on religion is clearly a part of the chapter on philosophical ethics which is called “*Ethica*.” Religion (*religio*) is Chapter A of general ethics (*Generalis*) in Baumgarten's text. This chapter A is followed by duties to oneself (B, *Officio erga te ipsum*), and duties toward others (C, *Officia erga alia*). The same division into two parts can also be found in the Menzer edition the Stark edition of Brauer. Nor do I understand why “*Ethica*” is translated as “Morality,” and not as “Ethics.” I shall use “*Ethica*” to refer to both “[On Religion]” and “On Morality.”

imperative. In the fourth and final part (Section 5), I will summarize the results and briefly evaluate the significance of my findings. Although I will take into consideration in each section how the Collins notes are related to (a) Baumgarten's textbook, (b) Kant's other published lectures, and (c) Kant's reflections on moral philosophy, the main focus will be on (d) how the main tenets of this lecture are related to Kant's mature work and why they are relevant for understanding Kant's philosophical development as well as his moral philosophy in general.

2 Ethics

Kant's late book on *The Metaphysics of Morals* contains, of course, two parts: the *Metaphysical First Principles of the Doctrine of Right* and the *Metaphysical First Principles of the Doctrine of Virtue*. The Collins lectures cover only the material of the second part of the *Metaphysics of Morals*. It is, however, surprising how closely the contents of the "*Ethica*" in the Collins notes correspond to the book that appeared in 1797, that is, more than twenty years later. The similarity is especially noticeable in the *Doctrine of the Elements of Duty* (MS 6:415–485). The "Introduction" of the *Metaphysical First Principles of the Doctrine of Virtue* which is meant to establish how the metaphysics of morals fits in with critical philosophy must, of course, be rather different from anything that can be found in lectures given many years before this critical philosophy was explicitly formulated. But there is one other fundamental difference. Whereas the first chapters of Baumgarten's *Ethica philosophica*, just like the first chapters of the "*Ethica*" in Collins, deal with religion, there is no such discussion in the *Metaphysics of Morals*. Instead, we find at the very end of the work a chapter entitled "Religion as the Doctrine of Duties to God Lies Beyond the Bounds of Pure Moral Philosophy" (MS 6:486–491). In it, he reaffirms not just that for him ethics is only about "the relations of human beings to human beings" and contains as such no duties to God (MS 6:491), but also that God enters into ethics only as an "idea" (MS 6:487). This differs almost as much from Baumgarten as it does from his own earlier view. In the Collins notes, God has a real role to play in morality. Kant seems to be convinced that "belief in a heavenly supplement to our incompleteness in morality make[s] up for our want" (C 27:317). We need "knowledge of God" to follow the moral law (C 27:315); and "natural religion should properly furnish the conclusion to ethics, and set the seal on morality" (C 27:305). If this were not enough, in the Collins notes Kant seems still quite confident in talking about "duties to God" (C 27:442), whereas he categorically rules out such a

notion in the *Metaphysics of Morals*.⁷ We will explore the reasons that this should be so in the next part of this chapter. For now, we will concentrate on the discussion of the duties of "morality proper" (C 27:339).

Just as in the *Metaphysics of Morals*, in the Collins lectures Kant divides up our duties into "duties to oneself" and "duties to others," roughly following Baumgarten's distinction. But in 1797, this distinction embeds another distinction, namely, that between perfect and imperfect duties. This is absent in Collins, but it can be found in Mrongovius (M II 29:609f). Apart from this nicety and some other minor differences such as the order of the treatment of particular duties, the discussion remains very much the same. Kant deals with duties to oneself under the broad categories "duties to ourselves as rational" and "as animal beings," although he discusses our duties as human animals in the Collins lectures only after discussing our duties as rational beings, whereas he discusses suicide, lust, and gluttony first in the *Metaphysics of Morals*.⁸ The reason for this change is probably that in the later text Kant adds another subdivision, namely, our duties toward ourselves as innate judges of ourselves, under which he discusses conscience (MS 6:439) and our presumed imperfect duties to improve ourselves naturally and morally (MS 6:442–444).⁹ Also, he renames the category of duties which we have to ourselves as "intelligent beings" to categories we have "merely as moral beings" and discusses just lying, avarice, and servility under this heading. In Collins, this part covers "proper self-esteem" (C 27:349–351), conscience (C 27:351–357), self-love (C 27:357–360), and self-mastery (C 27:360–369). We can see a greater emphasis on autonomy in the *Metaphysics of Morals* which shows itself especially in a section called "On the First Command of All Duties to Oneself," which commands the moral search for self-knowledge and contains the famous observation that "only the descent into self-knowledge can pave the way to holiness" (MS 6:441).¹⁰

⁷ Compare also MS 6:491 and C 27:307. But see also C 27:455, in which he rejects the idea that there can be a duty not to violate God's majesty. So, his view of "duties toward God" is already much narrower than that of many contemporary moralists.

⁸ The discussion of our duties as animal beings in Collins follows that of Baumgarten much more closely than the later text and is also much longer, containing some topics which are not explicitly discussed later. Altogether, it takes up thirty-five pages (C 27:369–413). In the *Metaphysics of Morals*, it is just seven pages long (MS 6:421–428). There are also frequent references to Baumgarten in Collins. Accordingly, the earlier discussion is much richer and should be used to complement the discussion of Kant's views on these subjects.

⁹ This section is prepared already in Collins, as Kant criticizes Baumgarten on his claim that the duties of self-perfection cover "all the [possible] perfections of man" (C 27:363).

¹⁰ I have substituted "self-knowledge" here for "self-cognition," as "self-cognition" seriously distorts Kant's meaning.

The discussion of our duties toward others in the *Metaphysics of Morals* is similarly dependent on the earlier discussion in Collins. But its rough division into (a) duties of good will and/or benevolence and (b) duties of indebtedness or rectitude is now replaced by a threefold division of “duties of love” – subdivided into duties of beneficence, gratitude, and sympathy (MS 6:452); duties we have because we owe respect to others (MS 6:462); and duties that depend on the conditions of others (MS 6:649). But this new division is more a reflection of Kant’s formal concerns than an indication of any change in Kant’s view about the substantive issues in question. Friendship remains important; contempt, ridicule, arrogance, and defamation remain just as much vices as in the earlier context. Again, the discussion in Collins is much more dependent on Baumgarten. It discusses many more phenomena than the *Metaphysics of Morals*. Yet the account of actual duties seems to have changed very little.

There is, however, one important material difference between Collins and the *Metaphysics of Morals* on duties, which concerns truthfulness or lying. Whereas in Collins these are discussed in the context of duties toward others, in the *Metaphysics of Morals* this discussion is set in the context of duties toward ourselves. In fact, lying is represented as “the greatest violation of a human being’s duty towards himself” as a moral being (MS 6:429). To say it again, in the Collins lectures, this discussion turns up in the context of duties toward others. One of Collins’ headings reads “Of Ethical Duties Towards Others, and especially Truthfulness” (C 27:444). This does seem to be a highly significant shift.

However, as the account of the moral duties we owe to others is so short in the *Metaphysics of Morals*, we cannot say much about this. It covers just twenty-six pages (MS 6:448–474) and is not even half as long as the part on self-regarding duties (MS 6:417–474, i.e., fifty-seven pages). In the Collins lectures, the account of duties to others takes up fifty-seven pages (C 27:413–470), whereas the discussion of duties toward ourselves is about seventy-three pages long (C 27:340–413). If we may take length as a rough approximation of what seemed important to Kant, it appears that although duties toward the self were always more important to Kant than duties toward others, they became even more important as he got older.¹¹ Compared to Baumgarten, we may in any case speak of a complete reversal, as his account of the duties toward others takes up about one hundred pages, whereas the duties we have to ourselves covers only fifty pages, not including Baumgarten’s duties to

¹¹ For the fundamental importance of duties to oneself, see also Louden 2000. The book is in many ways a defense of the centrality of duties to oneself in Kant.

animals and inanimate objects.¹² In other words, Kant moved further and further away from Baumgarten, making duties to ourselves more and more important, while eliminating not just duties to animals and inanimate objects but, ultimately, also duties toward God.

One of the most significant sections of the Collins lectures for Kant's later philosophy is perhaps the [last section](#), "Of the Final Destiny of the Human Race," which in some ways anticipates Kant's *Idea for a Universal History from a Cosmopolitan Point of View* of 1784. It also attests to the power of Kant's faith in education at this time. He answers the question as to whether we can justify any hope in the progress of humanity by pointing toward education:

The Basedow institutes of education create a small but fervent hope . . . Once human nature has attained to its full destiny and highest possible perfection, that will be the kingdom of God on earth, and inner conscience, justice, and equity will then hold sway, rather than the power of authority. This is the destined final end, and the highest moral perfection, to which the human race can attain, and for which, after the lapse of many centuries, we may still have hope. (C 27:471)¹³

It appears that later Kant looked for such signs not so much to education but to the American and French revolutions (see Kuehn [2001](#), 340–343).

3 Religion and morality

Let us now turn our attention to the part of the Collins lectures that are concerned with the subject matter of Baumgarten's *Initia* and contain Kant's early views on what later became the *Groundwork*. Kant's discussion here follows the outline of Baumgarten's *Initia* just as closely as Kant's discussion of ethics proper (Kant's "*Ethica*") follows that of Baumgarten's *Ethica*. Yet, even more than the *Ethica*, it also expands, modifies, and at the same time criticizes Baumgarten's theory extensively. Although Kant's "Proem" corresponds exactly to Baumgarten's *Prolegomena*, it is much longer. Where Baumgarten has nine short paragraphs on the nature of practical philosophy and universal practical philosophy, Kant extensively discusses the sense in which the term "practical" needs to be understood in moral philosophy. He emphasizes that practical philosophy is practical because its object is praxis,

¹² For which Kant severely criticizes him in Collins, saying "our author here commits an extravagance, in that he discusses duties towards inanimate things, animate but irrational creatures, and rational beings" (C 27:413).

¹³ See also Kuehn [2009b](#) and [2011](#).

not because it explicates its object in a practical way (C 27:245). Although it deals with the practical, it is essentially also a theoretical enterprise. More importantly, Kant also introduces in this context his view of the three kinds of imperatives of skill, prudence, and morality which will play such an important role in the second section of the *Groundwork* (G 4:414–417). Although the word “categorical” is missing, Kant already insists that moral imperatives hold without any conditions and must be understood as both objective and necessary (C 27:247). Yet he does not speak of “right” here but emphasizes in this context the “inner goodness” or “*innere Bonit  t*” of an action in accordance with what he calls the “moral imperative.” This seems to preclude the “priority of the right over the good” that many have observed in Kant’s mature moral philosophy.¹⁴

The “Proem” is followed by a chapter “On the Ethical Systems of Antiquity” which has no parallel in Baumgarten but seems to express one of Kant’s own concerns. It focuses on the *summum bonum* or on a topic that plays a large role in the second *Critique*. Yet here the *summum bonum* is discussed in the context of different ideals of the greatest moral perfection (C 27:252). Kant does not just focus on the ancient ideals but also discusses modern precedents to his view, such as those of Jean-Jacques Rousseau and David Hume. Whereas Rousseau, “that subtle Diogenes,” is aligned with the Cynic sect as someone who believes in the ideal of natural innocence, Hume is characterized as a follower of Epicurus and Zeno, who maintain, according to Kant, that morality is not a matter of nature, but of art (C 27:249). Plato’s “mystical ideal” is compared and contrasted with the “ideal of the New Testament (*Evangelium*)” which he considers the highest ideal and the greatest incentive for morality (C 27:251).

Kant claims that the ancient systems as well as the modern systems that follow them are essentially incomplete. The reason for this is that they cannot overcome human imperfection. In Christianity, by contrast, everything is complete. It possesses, according to Kant, the greatest purity and happiness in the highest degree, and it overcomes human imperfection with the help of God (C 27:251). Furthermore, Kant seems to believe in the earlier work that it is impossible even to become a moral person, or to “turn toward morality,” without a belief in God. This highlights again a highly significant change between the position of 1775 and 1785. Although religion and philosophy were still closely connected for the critical Kant, as the postulates of pure practical reason in the second *Critique* show quite clearly,

¹⁴ Rawls 1999, 27–30. We do not need to enter here into the controversy that arose around this claim.

he did not think that we need to believe in God to be moral. The categorical imperative is sufficient for doing what is moral, as we shall see.

But in Collins religion is indispensable. It gives morality more firmness, "beauty and reality" (C 27:307). Morality alone is insufficient to protect us. Indeed, Kant claims that "without religion all obligation is without motive. Religion is the condition for conceiving the binding force of the laws" (C 27:308). Without religion we would lack motivation for doing the moral thing:

All religion gives force, beauty and reality to morality, for in itself the latter is an ideal thing. If I imagine how beautiful it would be if all men were honest, such a state might induce me to be honest; but morality says thou shalt be moral as such and for thyself, be others what they may. The moral law then begins to become ideal in me; I am to follow the Idea of morality, without any hope of being happy, and this is impossible . . . So there must be one who gives force and reality to the moral law. (C 27:307)

That is, in other words, Kant believed around 1775 not just that morality needs the eudaimonistic component that he later completely abandons; he also believed that this eudaimonism is underwritten by God.¹⁵ God somehow grants that happiness that we need in order to be moral. Therefore, religion is of fundamental importance.

In other words, religion and morality are very closely related and depend on one another. Morality cannot exist without religion, while religion and theology cannot exist without morality (C 27:277).

Kant does not mean, however, "religion" in the sense of "institutional religion." He criticizes Baumgarten for suggesting that there is such a thing as "external" religion. Rather, religion for him is essentially internal. "True religion is that of the disposition" (C 27:308). The same is true of morality. Both are "a matter of the purest disposition," and these pure dispositions are such that "nobody except God can perceive them" (C 27:306f). That is the most important reason why we need to believe that God exists and why without God it would be very difficult, if not impossible, to engage in moral conduct. God – or, perhaps better, belief in God – is a necessary condition for the possibility of moral conduct. It provides a necessary aspect of moral motivation.

Still, it would be a mistake to say that even the early Kant tried to provide a religious foundation for morality. Kant clearly wants to say that morality is prior to religion and that religion is "morality applied to God" (C 27:305).

¹⁵ There are, of course, people who state that even the mature Kant introduces a eudaimonistic component into his moral theory through the notion of the highest good. I am not one of them.

Therefore, even natural religion, let alone theological religion, needs to be founded on morality. This critical view is already present in Collins. What have changed, however, in the critical works are the details of how religion is founded on morality.

Hand in hand with this change concerning the necessity of belief in God for motivational reasons goes Kant's reevaluation of religious or Christian morality. In the Collins lectures Kant essentially aligns himself with the Christian ideal, suggesting that it is also his own ideal. This will change radically in the *Groundwork*, in which the Christian ideal is considered inferior to the ideal of moral perfection. In fact, he even goes so far as to say that the Wolffian principle of perfection is better than the theological principle (G 4:443). And in the second *Critique* he explicitly says that the Christian ideal is just another material principle that is unfit to be the supreme moral law and is opposed to Kant's own formal practical principle of pure reason (KpV 5:41).

There is much more similarity and agreement between Collins and the critical works about the radical division between intellectual and empirical principles of morality and Kant's decisive rejection of any kind of empirical account. In both contexts Kant emphasizes the difference between empirical and intellectual or rational systems of morality, claiming that there cannot be any compromise or any mediated position between them. Those who accept empirical principles assume either internal sources for them or external ones. Internal sources may be either a moral feeling or moral sense or a physical feeling. Shaftesbury and Hutcheson are mentioned as representative of the advocates of moral feeling (C 27:253). Physical feelings were advocated in ancient times by Epicurus, in modern times by Helvetius and Mandeville (C 27:253). Philosophers who maintain that morality is based on external principles hold either that it is due to education, like Montaigne, or due to decrees of the government, like Hobbes. All these systems have a tendency to appeal to self-love as their main principle and they all tend to view morality as the result of accidental circumstances. As such, they are unsuitable as moral principles. They must be objective rather than subjective. They must also be rational.

It is therefore clear that by 1775 Kant has already decided in favor of a rational principle of morality. As the section "On the Ethical Systems of Antiquity" in Collins is clearly a precedent for the classification of all possible principles of morality based on heteronomy that can be found in the *Groundwork* and the second *Critique* (G 4:442ff and KpV 5:38–41), one might expect that Kant also distinguishes in the lectures between two rational principles, namely, the principle of perfection advocated by

Wolff, Baumgarten, and others, and the Christian principle that is characterized by a divine all-perfect will (G 4:443). But this is not what he does. Instead, he first lists three rational principles, namely, the principle of perfection of "Baron Wolff" which he considers "empty and unphilosophical"; the principle of truth put forward by Richard Cumberland (1631–1718), whom he takes to claim that "we all seek perfection . . . but are betrayed by illusion; morality, however, shows the truth" (C 27:277); and "Aristotle's principle of the mean" (C 27:276f). All three principles are, he claims, tautologous and therefore unsuitable for the supreme principle of morality. They may be rational, but they are empty. Kant then goes on to dismiss the Christian principle, or, perhaps better, the divine command theory, as inadequate as well. However, his reason is not that it is rational or tautologous, but that it is an external principle. The "pure intellectual principle must not . . . be a *principium externum*" (C 27:277). It cannot depend on the divine will that is outside of our own will, and to say that we should not lie, for instance, because it is forbidden by God would be a fundamental mistake because that would actually establish a one-sided dependence of morality on religion. It would make morality subservient to religion, and that would be just as wrong as making religion subservient to morality. "The question here is not whether theology is a motive to ethics, which it undoubtedly is, but whether the principle of moral appraisal is a theological one, and this it cannot be" (C 27:277).

Thus we may say that Kant counts himself together with Christianity among those who follow a rational principle and argues that morality is not in any way "an object of the senses, but an object merely of the understanding" (C 27:254), but we cannot say that morality simply consists in obeying commands given by God.¹⁶ The intellectual principle of morality is, rather, based on the "internal characteristics of the action insofar as we consider it by means of the understanding" (C 27:254).

The section "On the Highest Principle of Morality" therefore clearly spells out an argument that is meant to show that although the moral law is a command and thus might seem to imply someone who issues the command – that is, God – this is false. The moral law as a command *can* be understood as having been issued by God, but God is not its source. Put differently, Kant claims that the moral law is not the moral law simply because it comes from God, but rather that God himself chose the law in

¹⁶ It is not insignificant that Kant here uses "understanding" and not yet "reason," as this clearly indicates the mature distinction between these faculties has not yet been developed by him. He uses both in the sense of "intellect."

question because it is moral in itself. Kant is clearly aware of the so-called Euthyphro Dilemma, and his entire moral philosophy rests on the rejection of the idea that morality depends on the preferences of any being, be it mortal or be it divine. Morality is independent of any such preferences and should, in fact, determine the preferences of any being aware of the moral law. This also means, of course, that he rejects voluntarism in ethics outright.

Put differently again, it is for Kant a fundamental presupposition of ethics that as moral beings we are autonomous, and not passive or affected by any other force. If we were so affected in acting morally, morality could have only subjective or accidental force. But it “contains *objective laws of what we ought to do*, and *not of what we want to do*. It is not a species of inclination, but a caution against all inclination” (C 27:275). Morality “has no pathological principles” (C 27:275). Rather, “[m]orality is the conformity of the action to a universal law of free choice” (M I 27:1426). And we must conclude:

All morality is the relationship of the action to the universal rule. In all our actions, that which we call moral is according to rule – this is the essential part of morality, that our actions have the motivating ground in their rule. If I make it the basis for my actions that they conform to the universal rule which holds good at all times and for everyone, then they have arisen from the moral principle. (M I 27:1426f)¹⁷

The notion that morality is constituted by a “universal rule which holds good at all times and for everyone” is closely connected for Kant with our “maxims,” even though the word “maxim” seems to have a slightly different meaning in the Collins lecture, for the text of the lecture reads:

Every man who is against morality has his maxims. A precept (*Vorschrift*) is an objective law, by which we ought to act; but a maxim is a subjective law by which we actually do act. Everyone sees the moral law as something he can openly profess. But everyone sees his maxims as something that must be kept hidden, because they are contrary to morality, and cannot serve as a universal rule. (M I 27:1427)

If everyone had the goal to become wealthy, for instance, then a person seeking after wealth should keep quiet about it, as he could not succeed if he made this goal public. Because everyone else also tries to become wealthy, they would stand in his way and hinder him in reaching it. Not everyone can be wealthy.

¹⁷ It is important to note that this last formulation is not to be found in the text of the Collins notes. It has been supplied from Mrongovius I. The entire passage can, however, also be found in Kaehler 62–73.

Although it is not clear that Kant's conclusion about this maxim is right – modern capitalism appears to have proved him wrong – it is clear that his conception of maxim is very different from the one which he espouses in the *Groundwork*, in which a maxim is nothing but a “subjective principle of volition” (G 4:401 n.) or a “subjective principle of acting” (G 4:442 n.), having nothing to do with whether it can be espoused openly or not. There is nothing morally nefarious about a maxim in the *Groundwork*, and it does not have to be kept private. In fact, it is essential to the categorical imperative that some of the “subjective principles” can, in fact, be shown to be universal. Nor do the maxims have to be principles according to which we actually act. They could be merely entertained for the sake of considering a certain course of action. In any case, Kant seems to use the word “intention” for what he later calls “maxim,” and can thus claim: “I examine by the understanding whether the intention of the action is so constituted that it could be universal rule” (M I 27:1427).

How fundamental the difference between the earlier and the later view on maxims is would be an interesting question to discuss. Given constraints in length, however, I choose to concentrate on Kant's claim that morality is essentially expressed by objective rules and pay less attention to its connection with the “subjective principle of volition.” In both passages, maxims are contrasted with an “objective principle,” or “the practical law.” And it is no accident that this points in the direction of the categorical imperative as it is formulated in the *Groundwork*. Therefore, I will consider in the remainder of this chapter how far “the supreme principle of morality” of Collins can be identified with the formula of the categorical imperative found in Kant's later work. Is the earlier principle more or less identical with the categorical imperative, or is there a fundamental difference?

4 The categorical imperative

Kant's categorical imperative in the *Groundwork* of 1785, also called the universal imperative of duty, reads as follows: “Act as if the maxim of your action were to become by your will a universal law” (G 4:421). The supreme principle of morality is, as we have seen, connected with several concepts that are important for Kant's mature moral philosophy. Thus we find in both contexts the concepts of “universal law,” of “universalization,” of “nature,” of “subjective rules versus objective rules,” of “imperatives,” and of “maxims.” Even though Kant's concepts of “natural law” and “nature” changed considerably between 1775 and 1785, and even though the full effects of the so-called Copernican revolutions do not reveal themselves in

the Collins lectures, we may discount such differences as even less significant than the change in the meaning of “maxim.” Nor do we have to assume that “action” has changed a great deal for Kant. In fact, given what we have already said, it should be clear that for Kant any action worthy of the name must be determined by rational principles. So it might appear that the two principles are essentially identical.

This appearance is misleading, however. There is an essential difference between the earlier and the later contexts which has to do with Kant’s changed view of the relation of morality and religion which we have noticed several times already. In this context, it reveals itself in a distinction which is not to be found in the *Groundwork*. Right at the beginning of his discussion of the supreme principle of morality, Kant differentiates between what he calls the “*principium der dijudication der Verbindlichkeit*” and the “*principium der Execution oder Leistung der Verbindlichkeit*” or “*Richtschnur und Triebfeder*” (C 27:274). This is translated as the distinction between the “principle of appraisal of obligation” and the “principle of its performance or execution,” or between “guideline and motive.” I would have preferred “judgment” as a translation for “*dijudication*,” as the word usually is taken to mean “the act or process of judging or making a judicial decision or determination.” It has a legal background and can simply mean to make a judicial distinction. In other words, it suggests an active determination of what is right or wrong, not the classification of something or someone with regard to worth.¹⁸ What Kant is talking about is the determination of whether something is morally good or evil. We find in the text, “If the question is: what is morally good or not? . . . I judge the goodness or depravity of actions” according to laws (C 27:274). If I ask the question “what moves me to live according to this law? That is the principle of motive” (C 27:275). The first principle is objective, the second one is subjective; the first is an act of the understanding (or reason), the second is the “moral sense” or the “moral impulse.” Dijudication has to do with the head, motive with the heart. The former is not pathological, the latter is. In the former, we are “autonomous,” to use Kant’s later language; in the latter, we are affected by a third being, namely, God. And when we are motivated to act morally, “there must indeed be a third being that constrains us to do what is morally good. But for the making of moral judgments we have no need for any third being. All moral laws can be correct without such a being” (C 27:277f).

¹⁸ To “appraise” something is to assess the worth of something; this worth may be more or less. “Dijudication” suggests less of a gradual distinction and does not suggest utilitarianism in any way.

Kant claims that these two distinct principles have been confused and that because of this confusion all of morality has been misconstrued. He argues that it is important to realize that the supreme principle of morality can consist only in the *principium diiudicationis*, and that the *principium executionis* is not really part of *pure* practical philosophy.

The categorical imperative, by contrast, seems to be the combination of these two principles. It tells us *to act* only in accordance with that maxim which we have *judged* or *determined* to be a universal law. Indeed, we might say that its most important aspect has to do with moral motivation. We are to act “from duty,” not merely “in accordance with duty” (G 4:397). In other words, reason is what should motivate us. There has been much discussion regarding whether and how the universalizability of an action is sufficient to show that the maxim is moral or not.¹⁹ As Kant puts it in the “Introduction” of the *Metaphysics of Morals*: “Ethics does not give laws for *actions* (*jus* does that) but only for *maxims* of actions,” and as we all know maxims are, for the mature Kant, subjective “principles of volition” (MS 6:389 and G 4:401 n.). So the categorical imperative does not just determine what would be the right thing to do in a particular case, but it also provides the decisive test of whether the action has been done for the right reasons. This is also supported by the very first footnote of the *Doctrine of Virtue*, in which Kant claims that what is our duty in a particular case is “easily stated.” Even someone merely acquainted with moral philosophy can know what should be done. What is more important is to know “the inner principle of the will, namely, that consciousness of this duty be also the *incentive* to actions” (MS 6:375 n.; cf. Kuehn 2010).

Kant ascribes the position that it is sufficient for moral action to know one's duty to someone “versed in practical philosophy.” The true practical philosopher also knows the inner principle. As so often, Kant seems to have in mind himself in this criticism and his own doctrine of 1775 which involved the belief that it was the *principium diiudicationis* which was important, and that the inner principle of the will could well be left to a moral sense implanted by God, and therefore ultimately to a belief in God. Accordingly, what has changed between 1775 and 1785 is precisely that this cannot be if we are to develop a *pure* practical philosophy. The motivation or the *principium executionis* must for the mature Kant also be rational.

One might argue that Kant further pursues this point in his distinction between the “duty of virtue” (*Tugendpflicht*) which refers to an end “in

¹⁹ I will make no attempt to summarize the vast literature on this. But see O'Neill 1975 for a wide variety of objections.

itself” and is “at the same time” also a duty, on the one hand, and the many duties (*Tugendpflichten* or *officium ethicum s. virtutis*) we have, on the other. He also calls the *Tugendpflicht* the “ethical obligation.” And this ethical obligation is singular and concerns the “merely formal” aspect of moral obligation. It is contrasted with the many duties of virtue we also have (MS 6:383). The duty of virtue, also called *obligatio ethica*, precedes any particular duty which we may have. It has to do with our:

virtuous disposition (*Gesinnung*), the subjective determining ground to fulfill one’s duty, which extends to duties of right as well although they cannot, because of this, be called duties of virtue. – Hence all the *divisions* of ethics have to do only with duties of virtue. Viewed in terms of its formal principle, ethics is the science of how one is under obligation without any regard for any possible external lawgiving. (MS 6:410)

The categorical imperative as the supreme principle of morality essentially has reference to the “inner principle of the will” or to the principle of dutiful action. It is not just the principle of determining what duties we actually have. It is still the “supreme principle of morality” of the Collins lectures, that is, the *principium diiudicationis*, but it is also the “principle of the will” that makes for the moral worth of the action (G 4:400) which he talked about in *Groundwork*, which is a new version of the *principium executionis*, namely, one which has been cleansed from any of the merely empirical accretions of popular morality.

5 Lessons from Collins

If the preceding analysis is correct, we can see how difficult it was for Kant – or perhaps better, how long it took Kant – to free himself from theological morality. Although in the 1770s he had already rejected the idea that religion can dictate moral duties to human beings, it took him a long time to accept the view that religion has no say in motivating human beings to do what is moral. In fact, it may be argued that he ultimately never completely abandoned this view, and that the postulates of pure practical reason in the *Critique of Practical Reason* show his continued indebtedness to theology. His view that we can conceive of the “highest good” as an object of moral volition only on the assumption of God and the immortality of the soul, while we can at the same time act morally by following the categorical imperative (without the help of God), is indeed puzzling. Many philosophers have criticized Kant for the eudaimonist incursion into his moral philosophy and believe that it would have been better if he had

scrapped this part altogether. The present discussion has uncovered some reason for Kant's attachment to the doctrine that God is necessary for morality: this view belongs to the *Urgestein* of Kant's moral philosophy. It is not just a superficial attempt at pacifying the powers that be. Nor is it that Kant felt it necessary to reinvent God (for his servant Lampe) in the second *Critique* after he had put him to rest in the first. God was always there for him in the background. One might well speculate whether he would have disappeared altogether, if Kant had lived, say, another twenty years or so.

Yet, more important than these biographical speculations is another lesson that we can learn from contrasting Kant's mature position with his pre- or proto-critical position, namely, that moral motivation was more important to Kant than some of the traditional accounts suggest. In fact, it belonged to his earliest and to his deepest concerns. It appears to me that the investigation of how Kant's theory of moral motivation developed over the years may shed an interesting light on his final view. This observation does not mean, of course, that it will make it any more attractive to contemporary Kantians (even if they were to be swayed by historical considerations).

Mrongovius II: a supplement to the Groundwork of the Metaphysics of Morals

Jens Timmermann

There are many reasons why a reader of the Cambridge *Lectures on Ethics* volume will struggle to appreciate Moral Mrongovius II.¹ First, these are of course student notes, taken by Krzysztof Celestyn Mrongowiusz at the age of twenty, and even a gifted student is not always the most reliable guide to a lecturer's thoughts. Second, the set is inconsistent as well as incomplete. The text appears to be largely based on notes taken in class in the winter of 1784–1785, but in the course of preparing a clean copy Mrongovius initially sought to supplement them with passages from an earlier set of notes that belongs to the Collins/Kaehler family.² To make things worse, the manuscript does not cover the whole semester. It breaks off in mid-sentence. Third, Gerhard Lehmann's transcription of the notes, included in volume 29 of the Academy edition, is well and truly dreadful. There are on average about two or three significant transcription errors per printed page, often turning the meaning of Mrongovius' words on its head. Four entire sentences or clauses are missing because Lehmann was skipping manuscript lines.³

¹ The first, third, and fourth reasons are not specific to Mrongovius. They also apply to other lectures translated in the Cambridge volume.

² That is, Moral Mrongovius I. This concerns the following three passages: 29:601.5–11 (cf. M I, 27:1402.24–31), 29:601.12–33 (cf. M I, 27:1403.22–1404.1), and 29:603.23–27 (cf. M I, 27:1403.13–19). Moreover, there is an additional passage, hitherto unidentified, on page 59 of the manuscript (29:602). It is based on 27:1402.12–21 (the story of the simple-minded country girl).

³ Lehmann was working from a reprographic reproduction and never even went to see the Gdańsk manuscript. A random selection of errors: for "vorhergehen" (precede, M II 29:598.12) read "vorgehen" (take place); for "Einigkeit" (unity, 29:602.33) read "Reinigkeit" (purity); for "Fähigkeit" (capacity, 29:603.15) read "Rohigkeit" (rawness, barbarousness); for "Motiv" (motive, 29:605.30) read "subiectiv" (subjective); for "Lust" (pleasure, 29:608.17) read "Böse" (evil); for "Natur" (nature, 29:610.3) read "Maxime" (maxim); for "nirgend" (nowhere, 29:614.34) read "eigen" (peculiar); for "gibt aber das Recht ab" (gives up the right, 29:614.36f) read "giebt aber das nicht ab" (does not hand over); for "immer" (always, 29:618.33) read "niemals" (never); for "Beweis" (proof, 29:619.5) read "Hinderniß" (hindrance, obstacle); for "Tausend" (a thousand, 29:620.36) read "Taugenichte" (ne'er-do-wells); negation "nicht" (not) missing at 29:633.37; for "nur" (only, 29:634.3) read "neue" (new); for "sine" (without, 29:635.22) read "sive" (or, also called); and so forth. Sentences or clauses missing at M II 29:602.20, 29:612.20, 29:619.7 and 29:620.14. See Clemens Schwaiger's article on the

Fourth, Peter Heath translated little more than half of the Academy text.⁴ Finally, the quality of Heath's selective translation is poor.⁵

So, readers are likely to be grappling with an inaccurate partial translation of an error-ridden edition of patchy student notes from one of Kant's courses on moral philosophy. And yet, Moral Mrongovius II is an extraordinary document. It is the record of lectures Kant gave at the University of Königsberg when the manuscript of the *Groundwork of the Metaphysics of Morals* was being prepared for publication at Johann Friedrich Hartknoch's workshop in Riga. Departing from earlier practice, Kant rarely mentions the "author," Alexander Gottlieb Baumgarten. For the most part he clearly presents his own recent work, and many of the *Groundwork's* central themes – the good will, hypothetical and categorical imperatives, the central role of autonomy – are restated, sometimes more instructively than in the *Groundwork* itself. Moreover, Mrongovius' notes cover numerous topics that – like Kant's reflections on ancient philosophy, religion and the highest good, the philosophy of law, the nature of punishment – resurface only much later in the official published works, if they resurface at all.⁶

Academy edition of the lectures on ethics for further information about Lehmann's editorship in general (Schwaiger 2000), and Zelazny and Stark on Mrongovius' notes in particular (Zelazny and Stark 1987).

⁴ Some extremely interesting passages are missing, for example, the account of virtue and happiness at M II 29:600–602, the analysis of conflicts and types of obligation at M II 29:614, the discussion of merit and demerit with regard to God and humans at M II 29:631, and the last ten or so Academy pages, which contain discussions of the nature of moral legislation and of moral sanctions (M II 29:633–642).

⁵ There are many trivial mistakes, for example, confusing "Vernügen" (gratification) and "Vermögen" (Heath: "capacity," cf. M II 29:599.28) or the omission of a sentence at 29:632.14–15. Equally disconcertingly, Heath misconstrues phrases and expressions with which he is not familiar, for example, "come to the fore" for "vorkommen" (i.e., occur, 29:599.19f); "for if it did we should not have so many requirements" (cf. 29:610.7f) instead of "for without it [i.e. inclination] we would not have so many needs"; "if only it had already come about" (cf. 29:611.15) rather than "if only that [the wish to be virtuous] were sufficient to bring it about [virtue]"; "even though one might be aware of it" (cf. 29:620.12) should be rendered "even if it [the disposition] could be cognised" (but it cannot); at M II 29:612.4f, Heath makes Kant say that "such worth gives us a good will," whereas it is the good will that gives us worth (not vice versa). No discernable attempt was made to limit the number of different English words used for a single German word ("compulsion," "constraint," "coercion," "sanction" all used for "Zwang"; "dignity" and "worth" stand for "Würde," but "worth" is also used for "Wert"; "freedom" and "liberty" both translate "Freiheit"; "kingdom" as well as "realm" renders "Reich"; both "purpose" and "end" are used for "Zweck"; and so forth) or to use the standard terminology of the Cambridge edition ("reverence" rather than "respect" is used for "Achtung," but also for "Ehrfurcht," "man" is used to render "Mensch," etc.). If you are interested in what Mrongovius has to say, sell your copy of the *Lectures on Ethics* and spend the money on a ticket to Gdańsk. I shall be using my own translation of a provisionally revised German text throughout.

⁶ The comparison with Collins' notes is instructive. Collins apparently attended the same course of lectures on moral philosophy but got hold of an old set of lecture notes that does not reflect Kant's ethical thinking in the winter semester of 1784–1785. Collins is unaware of the formulation of the categorical imperative, the role of autonomy, and the one moral incentive called respect for the law; in Mrongovius II, as in the *Groundwork* and the *Critique of Practical Reason*, they take center stage.

I The good will as the supreme good

Mrongovius' discussion of the good will is an excellent example of the manuscript's potential for making the *Groundwork* more perspicuous. The notes help us understand the structure of the book, particularly the basic premise of Section I. The famous first sentence – a good will is the only thing conceivable that can be considered good without limitation (G 4:393) – is meant to be grounded in common human reason. That much is clear from the text.⁷ But we have to turn to Mrongovius to understand that the book's opening is a veiled reply of common human reason to a question raised by the ancients. It is a question that Kant, like Mill after him, still takes to define the subject matter of ethics, that of the highest good:

The ancient Greeks focused their efforts to determine the principle of morality on the question: What is the highest good? Among all that we call good most things are good conditionally, and nothing is good without limitation except the good will. (M II 29:599.23–26)

Mrongovius continues along familiar lines: understanding, strength and prudence are just conditionally good; in combination with a bad will they are extremely harmful. Even health, gratification, well-being and a cheerful heart are good only if someone enjoying them is a morally good person. He concludes:

A mere good will is therefore good merely without a limitation, just by itself, in every respect and under all circumstances. It is the only thing that is good without any other conditions, and yet it is not good completely. (M II 29:599.30–33)

Thus Mrongovius swiftly adds an important qualification, restated in the *Critique of Practical Reason* (KpV 5:110) but merely suggested by what Kant says in the *Groundwork*. The good will is not good completely, that is, it “does not constitute the whole good” (M II 29:599.34f). It is possible for something to be good unconditionally and yet to be incomplete, namely when other genuinely good things⁸ can be added to make it better:

But the possession of virtue is not yet the whole good. – Virtue is the greatest worth of a person, but one's condition must also be desirable. The greatest

⁷ Mrongovius goes so far as to assert: “Everyone knows that nothing in the world is absolutely good without limitation except a good will, and that this good will constitutes the limitation of everything – and that is why then it is good without limitation” (M II 29:607.15–17).

⁸ The highest good contains the only two things that are good intrinsically, virtue and happiness. To include instrumental goods would be “double counting” (and hence fallacious) because the things instrumental goods can be good for are already included in the highest good.

worth of one's condition is happiness. Virtue combined with happiness is therefore the highest good. (M II 29:599.39-600.3)

The distinction between the highest good in the sense of the supreme good and in the sense of the complete good is already in place in 1784.

Mrongovius then presents criticisms of Stoics and Epicureans. Their differences of opinion notwithstanding, both schools make the same mistake: they ignore the composite internal structure of the highest – that is, comprehensive – good by making virtue and happiness respectively the only good. In their own way, both schools thus fail to observe the radical distinction Kant detects between the two types of intrinsic value, moral and nonmoral. As in the *Groundwork* (G 4:390.20), a related charge is brought against Wolff, who defines duty as the necessity to act on the greatest balance of reasons (or “grounds”), regardless of their nature or origin (M II 29:598.6).⁹ Both the ancients and the moderns fail to appreciate moral commands for what they are: unconditional, fundamentally distinct and, in cases of conflict, rendering any other practical consideration normatively insignificant.

2 Imperatives and the value of action

Mrongovius introduces Kant's theory of imperatives in the usual manner: imperatives are either hypothetical, containing “the necessity of action as a means to ends,” or categorical, denoting “the practical necessity of action per se, without the motivating ground being contained in any other end” (M II 29:606.13–17).¹⁰ He proceeds to distinguish the two varieties of hypothetical imperative, and contrasts them with the categorical imperative:

The hypothetical imperative says which action is good either for any end whatsoever or for an end that is actual; whereas the categorical imperative says which action is good by itself. (M II 29:607.9–12)

It cannot be the case, Mrongovius tells us, that the morally good will acts to conform to a hypothetical imperative, “for then it would be good only in so far as the end is good, and no end is good without limitation” (M II 29:607.25f).

⁹ Wolff's terminology is reminiscent of Thomas Scanlon's theory of “primitive” practical reasons, which is susceptible to the same kind of Kantian objection.

¹⁰ “The problematic imperative occurs in all practical sciences, e.g. in geometry when I say: if you want to measure a tower, you must do it in this and that way. Those who do not want to measure the tower do not have to do it” (M II 29:606.22–26). This strongly suggests narrow scope interpretation of problematic imperatives. Morality cannot be based on such an imperative because “I do not want to do it” is not a permissible option.

As in the *Groundwork*, the categorical imperative is thus characterized as the principle that contains the law of an unconditionally good will (see G 4:437.5). An action is good unconditionally if and only if it is determined by the moral law.

But Mrongovius goes beyond the published account in that he spells out the reasons that underlie and support this striking axiological claim. This is important because time and again Kant's readers ask why in *Groundwork* I action not determined by the law – action from inclination, not from duty, if in conformity with it – is dismissed for a lack of moral value.

Mrongovius' answer is that only action determined by the moral law is good in a self-sufficient manner. Both moral and nonmoral actions are made right by the adequacy of the law on which the agent acts. Whether a nonmoral action is good depends on whether the result it is meant to realize is actually brought about.¹¹ There are actions, Mrongovius says some pages further into the manuscript, "whose goodness depends merely on the effect" (M II 29:610.12). These are actions according to hypothetical imperatives, which are guided by our perception of means-ends relations. The interest we take in such an action is merely mediate – we are actually interested in its result, regardless of how it is brought about. The action is willed merely as a means to the intended effect, and if the latter does not materialize the agent has done the wrong thing. The action was not good for the purpose it was intended to achieve. In itself the action is worthless, and its being rational at the time does nothing to change this.¹² The action is good just insofar as it produces the right result.¹³ But moral actions are not like that. They are "not in the least less good even if no effect whatsoever comes about" (M II 29:610.13f). Reason takes an immediate interest in the action; the effect – although the agent is whole-heartedly committed to it – is secondary. The mere act of moral willing has a distinctive kind of value, whereas all other kinds of willing "are good only just for the sake of the end" (M II 29:610.15f).

We have already seen that Kant does not intend to say that virtuous action is the only non-instrumental good. Kant was not a Stoic; one's own

¹¹ Like Mrongovius, I am setting aside the question whether the end is good, or whether the action is good not just *prima facie* (for a certain purpose) but all things considered.

¹² That is why an imaginary being that possesses only nonmoral practical reason is in no way elevated above nonrational animals. Instinct could have produced effective behavior, and may actually have been a superior guide (G 4:395.12–16).

¹³ This idea is at work in the *Groundwork*, at least implicitly. We take an immediate interest in the effect and only a mediate – or instrumental – interest in the action itself if we act on hypothetical imperatives, whereas moral judgment produces an immediate interest in the action (and only a mediate interest in the result), called respect for the law. Cf. the footnotes at G 4:413f and 4:459f, and Timmermann 2009b.

happiness is valued for its own sake in just the way that morality is. But of the two intrinsic goods it is the only one that directly pertains to action. The goodness of moral action is self-contained in a way that instrumental goodness is not. Mrongovius immediately proceeds to formulate the categorical imperative in terms of our own legislative powers, reminiscent of the *Groundwork*'s "kingdom of ends":

We must make all our resolutions thus, as if we were legislating with the maxim of our will. The human being sees himself in a system of rational beings as a legislating member, otherwise we are mere instruments. (M II 29:610.16–19)

It would seem that it is precisely our capacity to act in a way that is in itself valuable that elevates us above the class of entities that are good as means only.

3 Form and matter – how to make the categorical imperative work

Mrongovius' notes are also immensely helpful for understanding the formal nature of Kant's ethical theory. How can the principle, as he formulates it, "never to will anything unless I can also will that my maxim be at the same time a universal law" (M II 29:608.15f)¹⁴ yield concrete practical results? Since Georg August Tittel published his *Über Herrn Kant's Moralreform* in 1786, it has often been alleged that the categorical imperative is an "empty and sterile concept" (Tittel 1786, 33; "empty form, without materials," 88), and that it cannot be applied unless the empirical "consequences and effects" of an action are taken into account (33, cf. 14). John Stuart Mill, Georg Friedrich Hegel and many a clever undergraduate were to follow in Tittel's footsteps.

Kant's student appears to have understood the crucial point that his critics missed. Reason as such has nothing to say about consequences, whether envisaged, likely or actual.¹⁵ If pure practical reason were without a formal foundation and had only the assessment of consequences to base its

¹⁴ Mrongovius adds that then "the will is never in conflict with another" (*niemals andern zuwider*, M II 29:608.16f), presumably indicating that wills determined by the categorical imperative are in harmony with each other, as well as themselves (cf. KpV 5:28.4–28). He then defines evil – not "pleasure," as Lehmann and Heath would have it – as that which conflicts with reason (M II 29:608.17).

¹⁵ Reason approves of consequences indirectly, if they are the intended result of action on a purely formal principle. But whether an action is sanctioned by such a principle is precisely what is at stake here.

judgment on it would indeed be idle and empty. Consequences are agreeable or disagreeable to inclination, not to reason.

Why, then, does reason reject an action as immoral? Because the agent cannot consistently will the underlying maxim to be adopted by all agents at all times without being caught in a contradiction. An action is immoral because the agent is guilty of contradicting himself in action; and pure practical reason, like all modes of reason, approves of consistency and abhors contradictions.

So, how is an immoral agent guilty of being at odds with himself? Mrongovius clearly lays out the distinction between the two kinds of contradiction that can make an action immoral, usually called “contradiction in conception” and “contradiction in the will” in the modern literature on Kant’s ethics. An action is morally impossible either if “its maxim cannot take place as a universal law” or if “its maxim actually can, but we cannot will this” (M II 29:608.19–23 and, in very similar terms, 29:609.20–24). Two of the examples that follow are particularly instructive.

The first is that of theft, which violates a perfect duty to others. In its universalized form a maxim to steal when convenient would generate a contradiction:

If it were a universal rule to take from everyone what is his, mine and thine would cease entirely. For that which I would like to take from another a third party would take away from me. (M II 29:609.24–27)

As such, the problem with the agent’s larcenous maxim is not that, if universalized, the institution of property would cease to exist. There is nothing immediately self-contradictory about a world without personal property. The problem is that the thief commits the act to keep the stolen object as his own; and that a world without property in which the stolen object is at his sole, secure disposition is inconceivable. This constitutes a manifest contradiction. Stealing is immoral because the underlying maxim cannot be thought as a universal law. The consequences of a universally adopted maxim of larceny are taken into account, but they do not do the normative work required to dismiss the larcenous maxim.

The second example is that of charitable action, that is, a wide duty toward others. Mrongovius makes clear that the sense in which I would not be able to will a maxim of “lovelessness” (*Liebllosigkeit*, M II 29:609.27) is that it would thwart my own selfishness – the very attitude that motivates my own loveless course of action in the first place. If lack of charity were universally practiced, “then I would also suffer myself,” and this is

impossible for me to will twice over: I cannot will it *ex hypothesi*, because we assume for the sake of the thought experiment that I am in need. As Kant is not a Stoic, he takes human needs seriously – they are not something we can decide to switch off or ignore at will. But I am, second, particularly committed to realizing my urgent nonmoral ends as a loveless person who wants to keep his possessions for himself. Without a hint of consequentialism or emptiness, the natural selfishness of human beings in general and the affirmed selfishness of the uncharitable agent in particular join forces in an act of imaginative reasoning to show that the uncharitable maxim cannot without contradiction be willed as a universal law.

4 Autonomy

One of the great innovations of the lecture series on moral philosophy attended by Mrongovius is its emphasis on the autonomy of the human will. The pivotal concept of Kant's mature moral system is absent from all earlier sets of ethics notes. It is mentioned neither in Herder nor in Powalski, nor in the popular tradition of manuscripts exemplified by Brauer, Collins, Kaehler and Mrongovius' own earlier copy.¹⁶ It is only in the notes that are the subject of this chapter that autonomy is revealed as the foundation of Kant's ethical system.

Mrongovius derives the idea that the will must impose a moral law upon itself for there to be unconditionally valid moral commands from the notion of freedom, in much the same way as Kant does on the first page of *Groundwork* III. We learn that human beings are not "guided by nature." Unlike animals, they do not "receive" the laws of nature because they "are not determined by inclinations" (M II 29:630.13–15). Human action is not reducible to a mechanism set in motion to turn out the ends of nature. As we already know, when human beings act on prudential imperatives they are interested in the outcome, not in the action as such, and both the particular end and the causal law cognized by reason to direct the will are given by nature. But moral action is not like that, and yet there must be a law to determine its causality. Mrongovius concludes that "the will must be a law to itself" (M II 29:630.16f), and that freedom makes it the case that the will views itself as legislating. In sum, freedom is "the foundation of all morality" (M II 29:630.18).

¹⁶ There are, however, references to autonomy in the Feyerabend's lecture notes from Kant's 1784 lectures on natural law (e.g., F 22:1326.16).

Initially, however, the concept of autonomy is introduced at the end of a systematic discussion of five types of rival ethical theory and their principles, all declared to be inadequate:¹⁷

The principles mentioned previously are principles of heteronomy. This, however, is autonomy of the will in that the will can view itself as self-legislating in all actions. (M II 29:629.6–8)

My actions are good insofar as I “can consider my will to be self-legislating” when I act, which “gives my morality a high dignity” (M II 29:629.10–13). In what follows, Mrongovius contrasts the heteronomy of the kingdom of nature with the autonomy of a “kingdom of rational beings whose purpose is a universal system of ends.” In this kingdom “we consider ourselves as those who obey the law, but also as those who give the laws” (M II 29:629.18–21).¹⁸ Augustine and Leibniz are named as Kant’s sources.¹⁹

It is probably a result of the new emphasis on autonomy that Kant now rejects the possibility of active obligation – the “moral licence to obligate others” – which, following Baumgarten, he extensively focused on in earlier ethics lectures. “In actual fact,” Mrongovius writes, “there is only passive obligation” (M II 29:612.21):

I am necessitated solely by the moral law, and another does have the capacity to obligate but the obligator as such is an obligator merely through the moral law. (M II 29:612.21–24)

Curiously, this does not yet seem to apply to God.

5 Duties to God

The fact that the final type of theory rejected by Kant before turning to self-legislation is theological ethics suggests that autonomy originated as an

¹⁷ I cannot discuss Kant’s arguments in detail, but they are all significant philosophically in their own right. The principles are: (1) empirical external principles, for example, society and education; (2) the principle of happiness, that is, all forms of eudaimonism; (3) the principle of moral feeling; (4) the rational internal principle of perfection; (5) the external rational principle of theology, that is, divine command theory. This kind of overview is familiar from both the earlier lectures on ethics and from the published writings, but whereas the former do not mention the concept of autonomy the *Groundwork* and the second *Critique* introduce rival theories, under the dismissive heading of “heteronomy” or “material principles,” only once autonomy has been introduced (at G 4:441 and KpV 5:39, respectively).

¹⁸ Somewhat surprisingly, God is named as the “supreme legislator” immediately afterwards, apparently undermining the autonomy of the human will. I shall return to this topic in the subsequent section of this chapter.

¹⁹ That Leibniz’ “kingdom of grace” served as an inspiration or a model for the “kingdom of ends” is also apparent at M II 29:610.35: “In the kingdom of ends, a human being must regard himself as a legislating member, or of rational beings. Leibniz calls the kingdom of ends also moral principles of the kingdom of grace.” See also KrV B 840.

internalized version of divine command theory. The following passage, taken from the discussion of the “theological principle,” is even more explicit:

It seems as if duty had as its foundation the will of a legislator, not something one does according to one’s own will but according to the will of another. But the will of another is not the will of another being; but rather our own will, in so far we make it universal and consider it as a universal rule. (M II 29:627.13–17)

The problem with this attractive account of Kantian autonomy is that Mrongovius’ evidence is ambiguous. Kant does not yet seem to be prepared to argue that there can be no duties to God because God’s existence is not revealed to us, which is the argument of *Metaphysics of Morals* (MS 6:443f). There, we can still *view* the moral law as sanctioned by God as a matter of personal faith, in order to increase its motivational efficacy, but it is no more than a standpoint pure practical reason urges us to take up for the sake of moral practice. Mrongovius’ notes suggest that God is actually still in charge of giving the moral law, to the effect that there is a layer of duties to God in addition to duties to ourselves and to others, if – largely or completely? – coextensive with them. Let us inspect the argument in detail.

At first Mrongovius distinguishes between the author of the law and the author of the obligation, very much in the spirit of another well-known passage in the *Metaphysics of Morals* (MS 6:227):

He is not the legislator who is the author of the law; but he who is the author of the obligation of the law. The two can be different. (M II 29:633.27–29)

There is no author of the moral law, just as there is no author of the laws of mathematics. Yet there is an “author of the obligation of the law.” Bearing in mind what was said earlier about autonomy one would expect Mrongovius to assign that role to the human will. The law of our own will can then be expressed as the law of a divine legislator (cf. MS 6:227.16–18), but it is still the will’s own law. But that is not the route Kant takes in Mrongovius’ lecture notes. In 1784 God still seems to function as a genuine if supplementary moral legislator. Through God’s will “a new obligation is added” (M II 29:633.32), over and above the obligation that proceeds from our own autonomous will.²⁰ Mrongovius continues:

²⁰ He does not address the question how God can impose an obligation upon us within the epistemological restrictions of the *Critique of Pure Reason*.

He cannot be a legislator who does not at the same time have the good fortune within his control that pertains to following the law: only someone who also has the power to ensure that the law is being followed has legislative authority. God is not the author of morality, for then it would be due to his will and we would not be able to cognize it from nature either. It lies in the essence of things. In the same way, God is not through his will the author of the relation of mathematical figures. – But God's will gives rise to a new obligation if the happiness of his creatures is within his control and he uses it as a condition; thus he obligates everyone who wants to be happy to obey him. (M II 29:633.33–634.6)

According to the idea of divine legislation as the source of supplementary obligation there are two points of view from which we can consider ordinary moral duties, for example, a duty of beneficence: as a duty to the recipient and as a duty to God. Mrongovius draws on the dual nature of our obligations at several points, most strikingly perhaps when he says that all duties “are owed duties with regard to God.” With regard to human beings there can be meritorious duty because in our actions we can go beyond what we owe them. Then we do not “merely” do our duty, “as with duties of right,” but we impose a new duty of gratitude on those who benefit from our act of charity. But no merit can be acquired with regard to God. We cannot benefit God. With regard to him we are always just doing our duty (M II 29:631.23–28; and similarly 29:620.4–6, 29:636.4–9).²¹

At the same time, Kant denies that we possess knowledge of God's existence. Therefore, additional duties to God notwithstanding, we do not know that he will reward moral and punish immoral action. As in the scenario at the end of the Dialectic of the *Critique of Practical Reason*, Kant argues that virtue and happiness must be capable of coming apart, at least in principle. At M II 29:624 he says that if virtue were always followed by happiness “it would have a price.”

If God were merely kind the moral law would lack obligating force with regard to him. (M II 29:635.3f)

We would no longer be able to appreciate the unconditional character of moral commands.

God has not arranged things in the world to the effect that virtue is rewarded because then everything would be done from self-interest. (M II 29:637.18–20)

²¹ In fact, taking into account that there are juridical duties imposed by the legislature of the state, there are three tiers of legislation: autonomy, the civic commonwealth and the commands of God.

Kant needs the possibility of friction between the demands of morality and the prospect of happiness to demonstrate the unconditional nature of morality.²²

6 Religion as reassurance

First and foremost, the function of friction is educational. What clearly emerges from the earlier discussion of the worthiness of being happy is that a morally good person *ought* not to be unhappy. Still, the picture of the moral life that emerges from Kant's critical discussion of ethical eudaimonism is decidedly gloomy. In a memorable turn of phrase, Mrongovius says that "the consolation of the virtuous is the brevity of life" (M II 29:624.11f). Virtue does not make us happy. Rather, it

contributes much to the misfortune of human beings. Inner worth affords a human being consolation that does not let him drown entirely, but it is still not enjoyment . . . The virtuous man looks gloomy because of his punctiliousness; but the vicious man is cheerful. (M II 29:623.20–29)

Nor is it true that virtue is the best policy. For the most part, a "prudent rascal" (*ein kluger Spitzbube*) externally observes moral norms, but he departs from them when it is to his own manifest advantage (M II 29:623.30–32). Finally, the eudaimonist principle is

entirely at odds with the nature of morality. One must not be virtuous for the sake of advantage, or else all morality is lost entirely. Morality does not yield a profit – rather, it costs us. In the latter case it shines forth, for there virtue rests on its own inner worth. The rewards of virtue must not serve as its motivating grounds; not because we can do without all advantage, but because then it would be self-interest. Its inner worth must be its motivating ground. (M II 29:623.34–624.3).

Mrongovius' portrait of the gloomily punctilious agent does nothing to dispel the old accusation that Kant exaggerates the conflict between happiness and virtue, or the more modern fear that his rigorous account of duty demands too much of the moral agent. In fact, it may seem to lend credence to the traditional dour interpretation of Kantian ethics that Marcia Baron and others have been working so hard to discredit in recent years (see Baron 1995).

²² Kant returns to this topic repeatedly, most prominently in the [final chapter](#) of the *Dialectic* of the *Critique of Practical Reason*, 5:146–148; see Timmermann 2013a.

Yet there are two reasons why Kant is not guilty as charged. First, the pessimistic diagnosis we find in Mrongovius is obviously an empirical or descriptive thesis rather than a normative claim. An uncooperative world can give the impression that vice flourishes at the expense of virtue. But Kant emphasizes time and again that morally good people deserve to be happy, whereas bad people do not. He is not endorsing the virtuous agent's gloom; and whereas friction between happiness and virtue must be conceivable for the purpose of moral education, there need be no value in its being actual. Kant has simply come to the conclusion that a good person is unlikely to be happy in an unjust world.

The second point takes us back to what is special about Mrongovius' lecture notes. He is unique in making explicit how seriously Kant took the point that ultimately morality ought not to be bad for the virtuous agent, even as a challenge to his ethical theory: if it could be shown that virtue remains unrecognized, if it were certain that there was no reward for costly virtuous behavior, "that would clearly be a great objection against moral principles" (M II 29:637.16f).

If we are convinced that those who punctiliously do their duty deserve to be happy while at the same time we often see that they are not, there are in principle two ways of dealing with the problem. There is the response of the modern-day skeptic who uses cost to personal happiness as an excuse to reduce the requirements of duty. Kant acknowledges that human beings are susceptible to this line of reasoning, and that this would be an impediment to moral progress.

But there is another response, which is quite different from that of the skeptic. Kant is not prepared to let an unjust world lessen the stringency of moral requirements. He grasps the other horn of the dilemma to conclude that the virtuous agent's loss cannot be final. If as a matter of fact it is unrealistic to suppose that justice will be done in this world, it must be done in the next. It is thus the task of religion to *reassure* the person who struggles to be morally good that justice will be done:

Rewards must not be represented as motivating grounds, or else it is merely prudent conduct, but as confirmatory grounds for the correctness and truth of moral laws. Even if he lived in a world in which unhappiness is distributed in accordance with virtue, the most virtuous person would not lack motivating grounds but merely confirmatory grounds. Otherwise I do not know whether my moral science is a chimera (an ideal of the creative imagination). (M II 29:637.8–15)

In fact, as in the *Critique of Practical Reason* Kant accuses the Stoics and Epicureans of construing ethical theories that are unrealistic in their claim to

be felicitically self-sufficient by failing to acknowledge two heterogeneous values within the highest good. The mistake the Ancients made was to try

to squeeze the two pieces into one, when in fact they are quite different; for the one indicates the worth of the person, the other the worth of his condition . . . But there must be a tie; they could not otherwise be joined together. Now, this tie is religion. The ancients tried to do it without religion, but it did not work. Only through religion may I hope that someone who makes himself worthy of happiness can actually be happy. (M II 29:600.33–601.1)

Kant takes the worry of overdemandingness very seriously. The answer is religion.

7 Fragility and autocracy

It is worth bearing the distinction between grounds of motivation and grounds of confirmation in mind as we consider the question of how Mrongovius' Kant conceives of moral progress. The account of autonomy and heteronomy is followed by a discussion, along very familiar lines, of morality and legality, the letter and the spirit of the moral law:

Who merely acts according to legality acts according to the letter of the law. Respect for the law is the motivating ground of duty. One must not add pragmatic motivating grounds to the law because that makes hypocrites who just act according to the letter of the law and are not at all different from rascals with regard to their moral worth. (M II 29:629.34–39)

As at the end of Section I of the *Groundwork*, we hear about the fragility of human innocence, which cannot be relied upon to make us good. Innocence is desirable, “but it does not last.” It is “easily led astray, for it contains no lasting principle.” The desires of the agent continually grow, “and inadvertently innocence is left behind” (see M II 29:603.37–604.3). In Mrongovius' notes, however, we learn who the targets of these remarks are meant to be: Rousseau, who tried to retrieve and revive innocence, if in vain, and the Cynicism of Diogenes:

Diogenes, the head of the Cynic sect, said: one achieves [virtue] merely through the simplicity of nature. This was merely negative. If only one were not to corrupt the human being he would already be virtuous. (M II 29:600:20–23)

Kant's recommendation is different. A stronger medicine is needed, the pure and unmingled representation of the moral law. Again, there are overtones of the *Groundwork*. Morality must be presented “in its inner pure dignity,” the possibility of reward can at best be a secondary concern.

This method “would be more effective than now that everything is mixed together” (M II 29:624.9–11). Learning to act on pure incentives thus leads to autocracy:

When reason determines the will through the moral law it has the force of an incentive, and then it has not merely autonomy but also autocracy. Then it has both legislative and executive power. The autocracy of reason to determine the will according to the moral law would then be the moral feeling. A human being does actually possess sufficient force for this if only one teaches him see the strength and necessity of virtue. (M II 29:626.2–9)²³

8 Further perspectives

To conclude, let me mention just two further perspectives opened up by Mrongovius’ 1784 lecture notes on moral philosophy. First, toward the end of the manuscript there is an extensive discussion of rewards and punishments (M II 29:638–640). Mrongovius explores criminal punishment, as well as the rewards and punishments meted out by God. What is surprising is that Kant seems to have thought that an agent does not have to be morally good for God to make him happy. Not being vicious *may* just be good enough:

Under the rule of a supreme being those who are not entirely unworthy can be happy too. The necessary connection of happiness with the concept of a moral being in the idea of a perfect will is merit (*meritum*). Not unworthy and worthy are two different things. One is not unworthy when one is not liable to be punished. (M II 29:635.15–20)

Similarly, kindness is said to be “limited to the condition that the subject not make himself unworthy of it” (M II 29:641.14f).²⁴

Secondly, it is striking that in the winter of 1784 Kant had already worked out much of the philosophy of law as published in 1797, which is based on the concept of external freedom (see M II 29:618.30).²⁵ Consider, for example, the following passage, which is strongly reminiscent of the Doctrine of Right:

²³ By contrast, a few pages further into the text, while discussing God as the moral legislator, Mrongovius seems to revert to the view that God is needed to motivate: “The moral law gives obligation but not incentives. These are given to morality by religion. God is to be viewed as the ground of a special obligation because we can all demand happiness of him” (M II 29:634.38–635.1).

²⁴ Cf., possibly, KrV A 814/B 841, where the limit of happiness deserved is said to be one’s own *immoral* behavior.

²⁵ The main source for this is *Naturrecht Feyerabend*, dating from the summer of 1784.

Constraint can be combined with all juridical duties and one has a warrant to constrain another if one has a right on him. Constraint is a hindrance to freedom; if it is external it is an external hindrance of universal freedom. Resistance to a hindrance to universal freedom advances universal freedom and is therefore right. Now, wrong is a hindrance to universal freedom, constraint a hindrance to this hindrance or the removal of it; and thus an advancement of freedom. What advances universal freedom is right. (M II 29:631.12–20)

However, Mrongovius shows that there is at least one crucial difference. Rather neatly, “*honeste vive*” is declared to be the principle of ethics, “*neminem laede*” the principle of right in the state of nature, and “*suum cuique tribue*” the principle of right in the juridical state (see M II 29:631.31–33).²⁶ By the time Kant publishes the *Metaphysics of Morals* Ulpian’s “*honeste vive*” now tells us to be juridically honorable. It is no longer an ethical principle, but rather the first principle of duties of right, derived from the right of humanity within our own person (MS 6:236).

These examples serve to show that Mrongovius’ lecture notes on moral philosophy deserve to be edited, translated, and, above all, read with great care.²⁷

²⁶ Powalski (P 27:143f), Feyerabend (F 27:1336), and Vigilantius (V 27:527) also assign “*honeste vive*” to ethics, but they do not replicate Mrongovius’ spheres of responsibility.

²⁷ I owe thanks to Werner Stark (Marburg), Stefania Sychta (Gdańsk), Kate Moran (Brandeis), and to the editors for their patient help.

CHAPTER 5

Vigilantius: morality for humans

Robert B. Loudon

Above all, . . . you should have heard his lectures on ethics! Here Kant was no mere speculative philosopher, here he was also a spirited orator who enchanted heart and feeling just as much as he satisfied the understanding. Indeed, to listen to this pure and sublime doctrine of virtue from the mouth of its creator in person, delivered with such powerful philosophical eloquence, gave a heavenly delight. Oh, how often he moved us to tears, how often he forcibly shook our hearts, how often he raised our spirit and feeling from the fetters of selfish eudaimonism to the high self-consciousness of pure free will, to unconditional obedience to the law of reason, and to the exalted feeling of an unselfish fulfillment of duty! The immortal philosopher appeared to us then to be inspired with a heavenly power, and he also inspired us, who listened to him in complete amazement. His audience certainly never left a single lecture of his doctrine of morals without having become better human beings.

Reinhold Bernhard Jachmann, *Immanuel Kant geschildert in Briefen an einen Freund* (1804)¹

I Introduction: the importance of the Vigilantius lecture notes

Paul Menzer (1924b, 323) and Lewis White Beck (1963, ix) both open their discussions of Kant's lectures on ethics with this well-known remark from one of Kant's earliest biographers, and it is still a good way to begin. Even after we acknowledge that the above encomium was written "from the perspective of a student who uncritically adores his teacher" (Kuehn 2001, 12),² there remains much that rings true in Jachmann's remarks. For those who wish to understand Kant's notoriously difficult ethical theory, there are

¹ Translation mine, from Jachmann 1804, 30–31; cf. Groß 1912, 133–134.

² In his Preface, Jachmann confesses: "I am entirely taken with the greatness of the immortal man, to me he was everything" (translation mine; Jachmann 1804, X; reprinted in Groß 1912, 121).

several reasons why his classroom lectures on ethics are an invaluable resource. Kant was a popular teacher who took pride in his ability to connect with his audience. And partly because his audience “consisted largely of unsophisticated boys, younger than present-day college students, usually away from their rural homes for the first time, and for the most part ill-educated” (Schneewind 1997a, xvii), in his classroom lectures he often employs a simpler vocabulary and sentence structure than one finds in his published writings, offers more examples to illustrate his main points, and – particularly in the ethics lectures – tries harder to relate theory to practice in an effort to illuminate the moral life of human beings. John Macmurray, in his introduction to Louis Infield’s English translation of Menzer’s text, puts this last point well when he remarks that the ethics lectures

reveal Kant as a practical moralist, applying his mind earnestly to the detail of conduct, and supremely concerned with the social effects of his moral teaching, in a way that none of his more speculative writings can hope to do. From these lectures we can form a concrete idea of Kant’s own conception of the good life. (Macmurray 1930, xviii)

But what about the Vigilantius lecture notes, the focus of the present discussion? There are several additional important reasons for paying particular attention to this special set of notes from Kant’s moral philosophy course. Let’s begin with Vigilantius himself, the individual who transcribed Kant’s lecture. Johann Friedrich Vigilantius (1757–1823), a Königsberg lawyer and civil servant, was a personal friend of Kant’s who served as his legal advisor on a variety of matters – for example, the preparation of his will, “because” (as Kant himself remarked in a letter to Vigilantius) “I am a child in judicial matters” (Kant to Vigilantius, February 27, 1798, Br. 12:233).³ Vigilantius also belonged to Kant’s regular circle of dinner guests after 1787, was present when Kant died on February 12, 1804 (Vorländer 1924, 2:302), and was a pallbearer at his funeral sixteen days later. Indeed, in his “Instructions concerning his Burial,” written in 1799, Kant specifically asked whether “Herr Regierungsrat Vigilantius (or, in case of refusal, Herr Professor Rink)” would be so kind as to insure that “decent refreshments” were available to members of the funeral procession (Br. 12:417). And in a Supplement to his will written in 1799, Kant also requests that his “dear friend” Vigilantius be given the “golden medallion” awarded to Kant as a token of their relationship (Br. 12:417).

³ In citing Kant I have used the English translations of Kant’s writings in *The Cambridge Edition of the Works of Immanuel Kant* – with occasional modifications. Translations of Kantian texts not included in the Cambridge edition are my own.

Vigilantius also had more than a passing interest in Kant's philosophy, auditing not only the professor's ethics course in winter semester 1793–1794, but also physical geography (summer semester 1793), logic (summer semester 1793) and metaphysics (winter semester 1794–1795).⁴ Furthermore, Vigilantius was thirty-six years old and already an accomplished jurist when he audited Kant's ethics course in 1793–1794 – that is, unlike most of Kant's audience members, he was by no means an unsophisticated and inexperienced student. As Vorländer remarks, in describing Kant's circle of dinner companions, “and above all there was the civil servant, that is to say Councilor of the Higher Regional Court Vigilantius, who even as a grown man heard . . . all of Kant's lectures, and in addition to his broad education in particular was an excellent jurist” (1924, 2:301).⁵

But in addition to his professional accomplishments and strong personal relationship with Kant, Vigilantius also chose a particularly auspicious time to audit the professor's ethics course. Kant normally lectured on metaphysics during the winter semesters, but for the winter semester of 1793–1794 he decided to lecture on “Metaphysik der Sitten oder Allgemeine praktische Philosophie samt Ethik nach Baumgarten,” a departure that Arnoldt calls “very striking” (Arnoldt 1909, 5:174, cf. 321). This is the only semester that Kant chose the specific title “Metaphysik der Sitten” for his ethics lectures, and several commentators have surmised that “the cause of Kant's deviation from his normal routine was a desire to work through some of the themes later to appear in his 1797 publication of the same name” (Naragon 2009, 20; cf. Stark 2004b, 385). For instance, in a letter written to Erhard on December 21, 1792, Kant announces that he is “at work on [*unter Händen*

⁴ In his letter to Kant of September 18, 1793, Vigilantius thanks Kant for permission to sit in on his geography course (Br. 11:450f). The German text of *Metaphysik Vigilantius* is reprinted in AA 29:943–1040, and a complete English translation is available in Ameriks and Naragon 1997a, 417–506.

⁵ Paul Guyer writes: “The Vigilantius lectures are structured around the same distinction between duties of right and duties of virtue or ethical duties that divides the subsequent *Metaphysics of Morals*, but it is striking that in the lectures Kant goes over the duties of virtue twice, from V 27:541–586 and 600–712. One can only wonder whether Vigilantius did not produce his manuscript from several antecedent sources” (Guyer, this volume, note 3). Given Vigilantius's strong personal relationship to Kant, as well as his reputation as a leading citizen of Königsberg, I find this doubtful. If Vigilantius had produced his manuscript from several antecedent sources, why wouldn't he have indicated so? Similarly, Gerhard Lehmann, after recounting these and other details regarding Vigilantius's personal life, concludes that “there can be no doubt about the reliability of the contents” (Lehmann 1979, 1045) of this particular set of lecture notes. And Wilhelm Krauß, in his 1926 doctoral dissertation directed by Erich Adickes, examined fifteen different sets of lecture notes from Kant's moral philosophy classes, concluding that the Vigilantius lecture was “the most valuable of all those notes he had examined” (Krauß 1926, 94, 104; as cited by Naragon 2009, 20). Like Krauß, I believe that the Vigilantius notes are our single most trustworthy source of information for Kant's mature lectures on ethical theory.

habenden])” his “Metaphysics of Morals” (Br. 11:399), and perhaps⁶ he viewed the winter semester 1793–1794 as an opportune occasion “to test the content” of his forthcoming book in oral lectures (Stark 2004b, 385). At any rate, Vigilantius’s timing was serendipitous. The particular set of ethics lectures he transcribed in 1793–1794 gives us a unique preview of, as well as much needed help in understanding, the later *Metaphysics of Morals*.

Granted, the Vigilantius notes are not problem-free. The German text in AA 27 (and the corresponding English translation in Heath and Schneewind’s *Lectures on Ethics*) is merely a copy of Vigilantius’s original text, which has been lost. And we’re not even sure who produced the copy: it was “prepared by several different people of (yet) undetermined identity, but presumably associated with Rudolph Reicke or Emil Arnoldt and prepared in the late 1800s” (Naragon 2009, 18). And Gerhard Lehmann’s editing of the Vigilantius notes in AA 27 – like his editing of other lecture notes, including the Vigilantius metaphysics lectures from 1794–1795 in AA 29 – also leaves much to be desired.⁷ Nevertheless, despite these flaws, and for the multiple reasons indicated earlier, Vigilantius’s “Notes on the Lecture of Mr. Kant on the Metaphysics of Morals” (V 27:478) is an undervalued jewel – a jewel that not only shines by itself, but one which is also extremely useful and fruitful (cf. G 4:394).

2 The Vigilantius lecture notes and the *Metaphysics of Morals*: a very brief comparison

How does the general structure of Vigilantius’ lecture notes compare to that of the later *Metaphysics of Morals*? Surprisingly, even though the phrase “metaphysics of morals” occurs in the title of each text, and even though their composition dates vary by only about three years, the organization of these two texts is not at all the same. There are several reasons for this.

First, the “Baumgarten factor.” This is a set of lecture notes from one of Kant’s ethics courses, and Kant – in accordance with government regulations that were in effect at this time – used a textbook for the course. For

⁶ Unfortunately, “perhaps” is as far as we should go here, for Kant had also expressed to other correspondents his plans to publish a work on the metaphysics of morals at least twenty-seven years earlier. For instance, in a letter to Lambert of December 15, 1765, Kant claims that he has “already worked out” the contents of a little essay on the “Metaphysical Foundations of Practical Philosophy” (Br. 10:56). As Beck remarks, here Kant announces “a fateful plan that, in one form or another, was to attract, even almost to haunt, him throughout his mature life” (Beck 1960, 6).

⁷ See Naragon (2010, 19) for some specific criticisms of Lehmann’s editing of the Vigilantius ethics lecture, and Ameriks and Naragon 1997b, xxxviii–xli for an assessment of his editing of the Vigilantius metaphysics lecture.

most of his ethics courses, including the 1793–1794 winter semester course on the metaphysics of morals, Kant used two books by Alexander Gottlieb Baumgarten as texts: *Initia philosophiae practicae primae* (*Introduction to Practical First Philosophy*, third ed. 1760) and *Ethica Philosophica* (*Philosophical Ethics*, second ed. 1751, third ed. 1763). Although the issues he addresses in the course do not follow the same order of presentation that one finds in Baumgarten's books,⁸ Kant repeatedly refers to Baumgarten's views throughout his lectures – usually in order to indicate a disagreement. For instance, in *Vigilantius* § 21, Kant states: “In his practical philosophy § 39–46 Baumgarten⁹ has put forward various formulas which, as imperative, are supposed to serve for the general [*allgemeinen*] principle of all obligation, though Professor Kant rejects every one of them” (V 27:517; see also 503, 510, 539, 544, 546, 550, 607, 608, 622, 625, 662, 712). By contrast, in the *Metaphysics of Morals* Baumgarten is not explicitly referred to at all. In other words, in *Vigilantius* Kant develops his conception of a metaphysics of morals, at least in part, dialectically – by indicating his disagreements with Baumgarten's competing conception. Three years later, in the *Metaphysics of Morals*, the numerous disputes with Baumgarten are not mentioned at all.

Second, the “*Rechtslehre* factor.” Although both the *Vigilantius* notes and the *Metaphysics of Morals* are divided into paragraphs (§), *Vigilantius* (as one might expect in the case of a lecture transcription rather than a published book) contains no additional divisions into parts, chapters, sections, and so on, and has only a handful of headings.¹⁰ The *Metaphysics of Morals*, by contrast, is replete with numerous parts, chapters, and sections, and contains well over 150 different headings. But the two most fundamental and important parts of the *Metaphysics of Morals* – the *Rechtslehre* and the *Tugendlehre* – are not present in the basic structure of *Vigilantius*. Granted, there is a bit of “proto-*Rechtslehre*” in *Vigilantius* §§66–71 (27:587–600), where Kant discusses duties of right (*Rechtspflichten*) or *officia externa*, “for which external lawgiving is possible” (MS 6:239), and perhaps also in §§ 43–52 (V 27:551–557; cf. MS 6:331–337), where he discusses punishment,

⁸ For example, Baumgarten begins his *Philosophical Ethics* by discussing religion and duties to God (see Schneewind 1997a, xxiv), whereas this is the last topic discussed in *Vigilantius* (V 27:712–32). See also Ludwig's comparison between Baumgarten's *Initia* and *Vigilantius* in Ludwig 1988, 54f.

⁹ *Vigilantius* divides his text into 148 paragraphs (§). Baumgarten's *Ethica Philosophica* is divided into five hundred paragraphs. *Vigilantius*'s paragraph numbers do not refer to Baumgarten's paragraph numbers.

¹⁰ Six, to be precise, located at V 27:479, 488, 508, 630, 712, 729. But in some cases (e.g., 488) they don't seem terribly indicative of what follows.

and again in § 100 (V 27:637–642; cf. MS 6:277–280), where he touches on sex and marriage. But these short discussions pale in comparison with what we find later in the *Metaphysics of Morals*. For the *Rechtslehre* occupies well over half of the *Metaphysics of Morals* (169 out of 290 AA pages), whereas *Rechtslehre* topics occupy less than one-tenth of *Vigilantius* (24 out of 253 AA pages).

Why is there so much *Rechtslehre* material in the *Metaphysics of Morals*, and so little in the *Vigilantius* lecture notes? A combination of factors seems to contribute to this difference.¹¹ First, Baumgarten's *Ethica Philosophica* contains very little discussion of the topics that appear later in Kant's *Rechtslehre*.¹² And the conception of a metaphysics of morals articulated in the *Vigilantius* lecture, as noted earlier, is developed, at least in part, in opposition to Baumgarten's position. Second, the topics that are discussed later in the *Rechtslehre* (property, contract relations, natural vs. acquired rights, the power of the state, etc.) were topics that Kant typically taught in his Natural Right course. Kant taught a course on Natural Right twelve times between 1767 and 1788, and although he seems to have increasingly thought of *Rechtslehre* issues as properly forming part of practical philosophy (in large part because he views *Rechtslehre* issues as themselves being governed by a version of the categorical imperative), he was not explicit about the details of this matter prior to the publication of the *Metaphysics of Morals* in 1797. And third, some of Kant's most famous works in legal and political philosophy were published in the mid-1790s – for example, *On the Common Saying: That May Be Correct in Theory, but It Is of No Use in Practice* (1793), *Toward Perpetual Peace* (1795), as well as the second part of *The Conflict of the Faculties*, in which he argues that the French Revolution is a symbol of moral progress (AA 7:85f – written in 1795, published in 1798). Between 1793 (when *Vigilantius* first began taking notes in Kant's metaphysics of morals course) and 1797 (when the *Metaphysics of Morals* was finally published), Kant became increasingly interested in political issues.

Keeping both the Baumgarten and *Rechtslehre* factors in mind, here are two further observations concerning the basic structures of the *Vigilantius* lecture notes and the *Metaphysics of Morals*:

¹¹ I would like to thank Fred Rauscher for discussion on this issue.

¹² However, Baumgarten's *Initia* does have sections on Law, Punishment, and several other *Rechtslehre*-type topics. See Schneewind 1997a, xxiii. The complete Latin text of the 1760 edition Baumgarten's *Initia* is reprinted in AA 19:7–91. The 1751 and 1763 editions of his *Ethica Philosophica* are reprinted in AA 27:733–1028.

(1) Both texts conclude with discussions of religion and duties to God (V 27:712–732, MS 6:486–493). In *Vigilantius* this particular discussion is nearly four times as long as it is in the *Metaphysics of Morals*, but the content in both cases is similar. In each text Kant argues that “we can have no duties toward God” (V 27:713; cf. MS 6:487, 241, 443), but in each work he is also trying to carve a humanly necessary but qualified role for religion within ethics – to show that “*religion* is an integral part of the general *doctrine of duties*” (MS 6:487), but that religion is “founded [*gegründet*] on morality” (V 27:715) rather than vice versa.

(2) *Vigilantius* opens (§§ 1–63, 27:479–576) with a substantial discussion of the foundations of morality, or what Ludwig calls a “prolegomena to ethics” (Ludwig 1988, 54). Key topics discussed in this long opening part include the freedom of the human will, the dual nature of human beings, imperatives and the concept of duty, moral feeling, self-compulsion, and autonomy. Surprisingly, we don’t find anything comparable in the *Metaphysics of Morals*. As Schneewind remarks, “there is much more about the foundations of morality in §§ 1–63 [of *Vigilantius*] than there is anywhere in the *Metaphysics of Morals* (Schneewind 1997a, xviii).

In the remainder of my essay, I will focus my remarks on these two specific parts of the *Vigilantius* lecture notes (viz., the opening and concluding discussions). In each case, my goal is to draw attention to several important features within the specifically human dimension of this lecture. In the *Vigilantius* lecture, Kant – in part because of both the class lecture format and his audience, but also because of an important change in his thinking about ethical theory during the early 1790s – is working with a conception of ethical theory that is focused much more on human beings in particular (as opposed to rational beings as such) than we find in his better-known works on moral philosophy, including the later *Metaphysics of Morals*.

3 “The fragility of human nature”

In Kant’s opening discussion of the foundations of morality in §§ 1–63, we see at least three interrelated markers of the more human-oriented perspective that, for reasons outlined above, is on display in the *Vigilantius* lecture. First, the *Vigilantius* lecture represents a clear shift in Kant’s thinking about what a metaphysics of morals is – a shift away from viewing it as a body of a priori principles applicable to rational beings in general and toward one that, while still firmly committed to the claim that the supreme principle of morality is wholly a priori, sees its proper object as human beings. For

instance, at the beginning of the lecture, Kant states: “the metaphysics of morals is concerned especially [*besonders beschäftigt*] with the use of the freedom of the human will, according to rules of right [*Regeln des Rechts*]” (V 27:480).

This conception of a metaphysics of morals stands in sharp contrast to the one Kant employs earlier in the *Groundwork of the Metaphysics of Morals* (1785). In the *Groundwork*, Kant claims that the aim of “the metaphysics of morals is to investigate the idea and the principle of a possible *pure* will, and not the actions and conditions of human willing in general, which are largely drawn from psychology” (G 4:390). In the *Groundwork*, the project of a metaphysics of morals is “bound up [*verbunden*] with the concept of the will of a rational being in general [*überhaupt*]” (G 4:426), and it is purportedly “completely cleansed [*völlig gesäubert*] of everything that might in some way be empirical” (G 4:389, cf. 410).

In the Vigilantius lecture, by contrast, the conception of a metaphysics of morals Kant develops is one that is specifically bound up with the concept of the will of a human being, and this means that it cannot be completely cleansed of everything that might in some way be empirical. As others have noted (see, e.g., Wood 2002, 2–4), this more human-oriented perspective is also noticeable later in the *Metaphysics of Morals*. For instance, in the Introduction Kant warns readers that “a metaphysics of morals cannot dispense with principles of application, and we shall often have to take as our object the particular *nature* of the human being [*die besondere Natur des Menschen*], which is cognized only by experience” (MS 6:217). However, the human orientation of the Vigilantius lecture is arguably stronger than that of the later *Metaphysics of Morals*. In Vigilantius, the idea of a metaphysics of morals employed is one that is “concerned especially” with human wills. This implies that it is impure from the start, for the specific nature of human wills is something that can only be known by experience. By contrast, in the *Metaphysics of Morals* the focus is on a system of pure moral principles that is forced to take empirical information into account only when it seeks to apply these principles to human beings. This may appear to be merely an insignificant difference of degree only, but it is not. Indeed, Kant’s remark earlier in the Introduction to the *Metaphysics of Morals* that “concepts and judgments about ourselves and our deeds and omissions signify nothing moral at all [*gar nichts Sittliches*] if what they contain can be learned merely from experience” (MS 6:215) is not consistent with his earlier claim in the Vigilantius lecture that the metaphysics of morals is concerned especially with the use of the freedom of the human will. The Vigilantius lecture embraces a conception of the metaphysics of

morals that is impure from the start, while the *Metaphysics of Morals* brings in empirical information only at a later stage of application.¹³

A second sign of the specific human orientation of the Vigilantius lecture comes about ten pages later when Kant expands on his critique of Schiller's moral psychology, a discussion that he initiated earlier in 1793 with the publication of *Religion within the Boundaries of Reason Alone* (see RGV 6:23–4n.). At this second juncture, still more impurity enters into the picture. For here Kant is not just emphasizing the concept of a human will as opposed to a rational will in general. Rather, the discussion turns on a "subjective contingency" that "lies in the impulses of human nature [*in den Trieben der menschlichen Natur*]" (V 27:489). On Kant's view, it is necessary that these human *Triebe* oppose moral principle: "it is contrary to the nature of duty to *enjoy* [*gern*] having duties incumbent upon one; rather, it is necessary [*nothwendig*] that the impulses of the human being should make him disinclined to fulfill the moral laws, and that these impulses should be overcome only through the authority of the latter" (V 27:490). Schiller, by contrast, in his 1793 essay *Über Anmut und Würde* ("On Grace and Dignity") and elsewhere, holds that "in Kantian moral philosophy the idea of duty is represented with a severity that frightens away all the graces," and that a better model of human morality is one where impulses and laws, sensibility and reason, are completely in harmony with one another "and the human being is at one with himself" (Schiller 1966, 171, 167).

These are very different moral psychologies. Kant's position is that if humans were not subject to these impulses of nature, then there would be no moral imperatives – we would follow moral laws without necessitation. "The human being can . . . never attain to holiness in his being . . . But were he a holy being, he would have no motive for transgressing the moral law; he would have no duty, for want of necessitation" (V 27:492; cf. G 4:414). But he holds that the antecedent stated a moment ago ("if human beings were not subject to these impulses of nature . . .") is physically impossible (contrary to human nature) and would literally transform humans into God – "a being that is *alone holy*, i.e., has the property that He follows moral laws without necessitation" (V 27:489). Schiller, on the other hand, does not think that all humans necessarily remain saddled with *Triebe* that must be constrained by reason in order to be effective moral agents. In the case of the "beautiful soul" who represents "the symbol of complete humanity" (Schiller 1966, 173), there exists a full integration between *Triebe* and *Gesetze*, impulses and moral laws, one that enables him to act with a grace

¹³ For related discussion, see Loudon 2000, 3–30.

that is missing in the Kantian dutiful agent. The Schillerian beautiful soul, unlike the Kantian dutiful agent, no longer experiences any conflict between duty and desire, and this ideal of complete psychic harmony is one that Schiller urges all humans to aim at. For in this person, morality “*has become nature*” (Schiller 1966, 170) and there is no longer any need to constantly consult reason in deciding how to act. Rather, the beautiful soul can “with a certain security trust in the voice of impulse . . . without fear of being misled by it” (Schiller 1966, 173).

Kant agrees with Schiller that genuinely virtuous human beings act with joy (see RGV 6:24n.) and “good cheer [*mit frohem Muth*]” (V 27:491), but for Kant this does not mean that necessitation is absent in their moral conduct: “Where there is no necessitation, there is also no moral imperative, no obligation, duty, virtue, ought or constraint is conceivable. Hence the moral laws are also called *laws of duty* because they presuppose an agent subject to the impulses of nature” (V 27:489). And toward the end of his critique of Schiller’s position, Kant may appear to grant a concession to his opponent when he concedes that “it would be good if human beings were so perfect that they fulfilled their duties from a free impulse [*aus freiem Triebe*], without constraint and laws [*ohne Zwang und Gesetze*]; but this is beyond the horizon of human nature” (V 27:491). However, unlike Schiller, Kant does not think it is humanly possible to go beyond this horizon. To be human means to have a “moral disposition *in conflict*” (KpV 5:84), and the total psychic harmony advocated by Schiller is simply an illusion. To suppose otherwise is to fall victim to “*moral Schwärmerei*” – “an overstepping of the boundaries that practical pure reason sets to humanity” (KpV 5:85).¹⁴

Finally, and partially as a result of this, further on in Vigilantius we also find a distinctively Kantian conception of moral virtue that is tied to the realistic portrait of human nature articulated earlier. In § 59 Kant states: “virtue consists precisely in the strength of the resolve to fulfill our duties, and to strive against the constant enticements to do otherwise which sensible [*sinnliche*] feelings inspire” (V 27:570). On Kant’s view, virtue is a humans-only concept, applicable to us alone owing to our dual nature as both sensible and intelligible beings: “in him [*viz.*, the human being] . . . is virtue alone thinkable, for only where necessitation is the ground can one assume a steadfast determination in obeying the moral law” (V 27:490, see also 504f). However, the strength of resolve to listen consistently to reason

¹⁴ I am indebted here to Baxley’s insightful discussion of Schiller’s position as well as Kant’s response to it in Vigilantius. See Baxley 2010, 87–97, 111–115.

and thereby fulfill one's duties as part of one's moral character that is denoted by Kantian virtue should by no means be confused or conflated with the suppression of desire – a point Kant also makes clear in *Vigilantius*: “Now if one had to suppose here an overpowering or predominance, with total suppression [*mit völliger Unterdrückung*], there would no longer be any virtue at all” (V 27:570). In order for a creature to be human, both desire and reason must be present, for “the human being is no being sufficient unto himself; rather he depends on needs, and this is what the ground and efficacy of his sensible impulses actually consists of” (V 27:520). The specific strength of will tied to Kantian moral virtue implies not a total suppression of desire (if we had no desires we would not be human) but rather a balance between reason and desire that gives the agent self-control and which constitutes “the state of *health* proper to a human being” (MS 6:384). “The true strength of virtue is *the mind at rest* [*das Gemüth in Ruhe*] with a considered and firm resolution to put the law of virtue into practice. That is the state of *health* in the moral life” (MS 6:409; cf. V 27:492).

In the *Vigilantius* lecture, Kant expresses an acute awareness of “the fragility of human nature,” a fragility that stems from our “propensity to evil” (V 27:571; see also RGV 6:19–53) – that is, our universal tendency to place our inclinations and impulses above the moral law. And because of this fragility, morality will always remain difficult for us, in large part because our own self-conceit blinds us from the problem in the first place (see G 4:407, RGV 6:36). To suppose that we can permanently escape from this fragility and become “beautiful souls” at one with ourselves is, again, moral *Schwärmerei*. Rather, the best that humans can hope for is to acquire the vigilance and strength of will that is signified by virtue.¹⁵

4 “A necessary accompaniment to human nature”

The *Vigilantius* lecture and the *Metaphysics of Morals* both end with discussions of religion and duties to God, and, as remarked earlier, this seems to be one of the most obvious similarities between the two texts.¹⁶ But not all commentators acknowledge a similarity. Schneewind, for instance, in comparing the two texts, observes: “the lectures conclude, in §§ 138–148, with an examination of our duties to God. But the book ends with a section

¹⁵ For related discussion, see Loudon 2011, 16–24, 107–120.

¹⁶ It is also worth pointing out that none of the other major ethics lectures – Herder, Collins, Powalski, Mrongovius II – end on this theme.

strikingly entitled ‘Religion as the Doctrine of Duties to God Lies beyond the Bounds of Pure Moral Philosophy’” (Schneewind 1997a, xviii).

However, a closer look at the actual content of each discussion does reveal a similarity in position. Granted, the heading “Duties to [*gegen*] God” (V 27:712) appears right above § 138 in Vigilantius, but midway through § 138 Kant states clearly that “we can have no duties to God” (V 27:713), and this is also the same position he defends later in the *Metaphysics of Morals*. We have a duty “*with regard to* [*in Ansehung*]” God, but not a duty “*to* God” (MS 6:443, 487). His basic argument in both texts is that we can have duties only to rational beings that are objects of experience, and God does not fulfill the latter necessary condition (see V 27:713, MS 6:443f).¹⁷ The ground of our knowledge of God “is drawn neither from experience alone, nor from reason alone, and is equally little dependent solely on a mere insight as to the nature of an actual deity; it is based, rather, on grounds which are the source of the moral actions of human beings, and is thus practical” (V 27:724, see also MS 6:443f). Our idea of God is practical rather than theoretical, but nevertheless it is at the same time “a duty of the human being to himself to apply this idea, which presents itself unavoidably to reason, to the moral law in us, where it is of the greatest moral fruitfulness. In this (practical) sense it can therefore be said that to have religion is a duty of the human being to himself” (MS 6:444).

In other words, what is popularly referred to as a duty to God is properly understood as a human being’s duty to himself – a duty on the part of humans to have faith in the necessary conditions for the actualization of the moral law, a law that reason shows us is categorically necessary. And it is a duty that applies specifically to human beings, but not necessarily to other rational beings – the ground for this duty is “only subjectively logical” (MS 6:487) as opposed to objectively logical. As one commentator observes, this means that it is a duty “exclusively caused by the peculiarity of our faculty of cognition” (Ricken 2013, 418). From the perspective of a pure and a priori metaphysics of morals, there are no duties to God. But from the impure (empirical) perspective of human beings, we have duties with regard to God. “As far as mere reason can judge, a human being has duties only to

¹⁷ VMS 23:207–417 contains some early drafts of MS, and the position Kant advocates here regarding duty and God is also similar to what we find at the end of both V and MS. For instance, at VMS 23:416 Kant writes: “There can be duties *with regard to* [*in Ansehung*] certain merely possible beings which are nevertheless not duties *to* [*gegen*] these beings.” And he then proceeds to argue that the former are properly construed as “duties to ourselves” (VMS 23:416), a strategy also adopted in V and MS.

human beings (himself and others)” (MS 6:442)¹⁸ – this is why Kant entitles the Conclusion to the *Metaphysics of Morals* “Religion as the Doctrine of Duties to God Lies beyond the Limits of Pure Moral Philosophy” (MS 6:486). But Kant’s “duties to humans only” doctrine is by no means intended to close off the humanly necessary help of religious faith in pursuing moral aims that reason deems categorically necessary. Again, humans have a duty to themselves to view morality from the perspective of practical faith. As Kant puts it in *Vigilantius*: “only human duties are thinkable [*denkbar*], to oneself and others, though the duties themselves are conceived [*gedacht*] under the image of divine commands” (V 27:713). We know the moral law and what it requires of us through reason alone, but it is nevertheless “a necessary accompaniment [*eine nothwendige Begleitung*] to human nature” (V 27:530) to view our moral duties as divine commands. Without this allegedly necessary accompaniment, humans – once they realize that the prospects of achieving what morality demands are nil without divine assistance – will fall into despair and slack off in their moral obligations. Their commitment to morality will no longer be “rationally stable” (see Hare 2009).

The specific background of the remarks about the connection between morality and religious faith at the end of *Vigilantius* is Kant’s famous argument in the Preface of *Religion Within the Boundaries of Mere Reason* that “morality inevitably leads to religion” (RGV 6:6), along with his related discussion of our duty to promote the highest good, the best known version of which is presented in the Dialectic of the *Critique of Practical Reason*. The broader background is the effort of many, many Enlightenment intellectuals to redirect religion away from theological and toward moral purposes. “Religion is . . . founded on morality” (V 27:715), Kant declares in *Vigilantius*, and this core conviction is shared by Lessing, Rousseau, Hume (or at least Hume’s Cleanthes), Paine, Franklin, Jefferson, and a host of others.¹⁹ But as a corollary to our earlier argument concerning the specifically human perspective adopted by Kant in his classroom lectures on ethics, in *Vigilantius* we find Kant discussing religious faith specifically from the standpoint of human nature. Given human nature, humans need to view morality “under the image of divine commands” (V 27:713). Let me briefly sketch some of the relevant background points from *Religion* and the

¹⁸ In recent years Kant’s “duties to humans only” doctrine has been challenged repeatedly from the animal side, sometimes even by Kantians, see, for example, Korsgaard 2011. For a more sympathetic account of Kant’s position, see Kain 2010. But to the best of my knowledge, it has not been challenged from the God side.

¹⁹ For discussion and references, see Louden 2007b, 15–25.

second *Critique*, and then return to Kant's manner of addressing these issues in *Vigilantius*.

In the Preface to *Religion*, Kant starts out by saying that "morality in no way needs religion" (RGV 6:3), but within a few pages he shifts to the seemingly inconsistent claim that "morality thus inevitably leads to religion" (RGV 6:6). As Leslie Stevenson notes,

the thought behind the first sentence is that we can know what our moral duties are without any appeal to any idea of God, and that we can and should be motivated to fulfill them purely by the thought that they are indeed our duties (and not by any idea of divine rewards or punishments). In contrast, Kant's thought behind the second sentence is that we need to be able to *hope* that doing our duties will in the long run tend to promote "the highest good" – that is, the happiness of human beings in proportion to their virtue – and that this can only be brought about by God. (Stevenson 2006, 214)

In his discussions of the highest good, Kant often claims that "the ancient philosophers . . . in their investigation and research into ethics, reduced everything to the question, What is the *summum bonum*?" (V 27:482; cf. P 27:101; C 27:247; M II 29:599). Kant disagrees with the ancients on at least two points here. First, he thinks they are mistaken in reducing everything to the highest good. In the *Critique of Practical Reason*, for instance, he criticizes the ancients for trying to make the highest good the criterion of morally right action, "whereas they should have first searched for a law that determined the will a priori and immediately" (KpV 5:64). On Kant's view, ethical inquiry should begin not with the highest good, but rather with the moral law. The highest good grows out of the moral law and not vice versa (see RGV 6:5). Second, Kant's own conception of the highest good also differs substantially from the ancients. He defines the highest good tersely as "happiness distributed in exact proportion to morality" (KpV 5:110), and he envisions this occurring on a universal scale.²⁰ The highest good is realized when happiness is distributed universally among all moral agents in exact proportion to their virtue. But this gargantuan distribution task – a task that must also be carried out "disinterestedly" and "solely from impartial reason" (KpV 5:124) if it is to be done correctly – is far beyond the capacities of finite human beings. For (among other things) human beings do not even know who among them really is virtuous (see G 4:407; RGV 6:51; MS 6:392) – our own moral status (as well as that of others) remains inscrutable to us.

²⁰ Most of the ancients did not conceive the highest good in this manner. In the case of Aristotle, for instance, achieving the highest good – *eudaimonia* – does not involve either thinking or acting globally. For some additional differences, see Loudon (in press).

Nevertheless, reason shows us that we have a categorical obligation to promote the highest good – indeed, “we are determined *a priori* by reason to promote [it] with all our powers” (KU 5:453; cf. KpV 5:114; RGV 6:5). Reason tells us that we are obligated to promote this goal with all of our powers, and yet we quickly realize that our own powers are woefully inadequate to the task that reason demands of us. We are therefore also obligated to have faith in those necessary conditions that would make the realization of the highest good possible – one of which is the supposition that there exists an omniscient, omnipotent, omnibenevolent being who can accurately apportion virtue to happiness on a universal scale. Our reason thus finds the possibility of the highest good thinkable only on the presupposition of a supreme intelligence, and to assume the existence of this supreme intelligence is itself connected with the consciousness of our duty (see KpV 5:125).

Kant’s explanation of the need for faith in God at the end of *Vigilantius* squares with his earlier arguments presented in the second *Critique* and in *Religion*, but if one looks carefully one can also find additional pleas concerning why the special situation of humans makes this faith necessary. Humans are brought to “the supposition of theism,” he notes, “solely through morality and not through physico-teleology,” (V 27:717), and the need for the supposition “lies merely in the moral need for us [*liegt bloß in dem moralischen Bedürfniß für uns*] of such a being” (V 27:718). “The human being assumes [*der Mensch nimmt . . . an*] as a practical hypothesis, that if it is true that the grounds of morality and its laws lead only to the highest end, a supreme being must actually exist, whereby the highest good is possible, in that otherwise the highest end of his moral actions would not be possible. The concept of God therefore presupposes morality . . .” (V 27:724).

The Kantian faith in God is thus a human assumption “based on grounds which are the source of moral actions of human beings” (V 27:724); it lies “merely in the moral need for us” (V 27:718), and it is “a necessary accompaniment to human nature” (V 27:530). So would the grounds for such a faith be absent in the case of other rational beings that weren’t human? Would a different, nonhuman us not have this need, not make this assumption? And would the supposition of God not be a necessary accompaniment to rational beings that were not humans? Surely what Kant thinks he is defending in the *Vigilantius* lecture is a more traditional conception of faith – one in which “faith is still understood as a taking to be true of what actually is” (Ameriks 2012, 252). But in his frequent appeals to human moral needs as the ground for the supposition of theism, one wonders whether he is in fact defending something quite different. Is Kant’s God simply a

proto-Feuerbachian projection of human needs? For Feuerbach too will argue later that “God springs out of the *feeling of a want*, what the human being is *in need of*” (Feuerbach 1971, 134). This is not where Kant wants to go, but the trajectory of his thought may have inadvertently opened the way for a quite different “true, i.e. anthropological, essence of religion” (Feuerbach 1971, 3).

In the Vigilantius lecture, Kantian ethics comes down to earth to examine the fragility of human nature and to propose a necessary accompaniment to this fragile nature. Reading these *Notes on the Lecture of Mr. Kant on the Metaphysics of Morals* (V 27:478) might not necessarily make us “better human beings” (Jachmann, reprinted in Groß 1912, 134), but they do reveal much about the development of his mature moral theory, and they also prove that “a Metaphysics of Morals, regardless of its daunting title, is still capable of a great degree of popularity and suitability for the common understanding” (G 4:391).²¹

²¹ An earlier version of this essay was presented as an invited paper at the VIII Kant Colloquium, Marília, Brazil, and also at the Universidade Federal do Rio Grande do Sul in Porto Alegre, Brazil and the Universidade Federal de Pernambuco in Recife, Brazil, in August 2013. I am very grateful for the opportunity to present my work in Brazil, and for helpful comments raised during discussion.

PART II

Practical Philosophy

CHAPTER 6

Ancient insights in Kant's conception of the highest good

Stephen Engstrom

Kant's moral philosophy has long been depicted, by detractors and defenders alike, as opposed to teleological approaches in ethics, and in particular to the idea, associated with the schools of ancient Greek ethics, that *eudaimonia* is the highest practical good and ultimate end in human life. There is plainly an element of truth in this representation, to the extent that *eudaimonia*, as understood by at least some ancient authors, can be identified with happiness as Kant understands it (*Glückseligkeit*). Moral knowledge, he argues, is distinguished from other modes of practical thought by an unconditioned necessity that marks it out as resting on an a priori principle, a formal condition of the possibility of knowledge of the good, requiring that the principles implicit in the practical judgments in which we determine how we should act could be universally agreed on as practical laws. It follows that moral knowledge cannot lie in prudence, the practical intelligence that enables the generally effective pursuit of one's happiness. Since one person's happiness differs from another's, prudence cannot by itself yield the universal agreement requisite for knowledge of the good, an agreement that ensures harmony in action. Philosophical accounts of morality that take happiness to be the ultimate good are therefore, he holds, to be rejected.¹

Kant's opposition to such accounts, however, has often been thought to reflect an opposition to happiness itself. In his two most famous works in practical philosophy – the *Groundwork of the Metaphysics of Morals* (1785) and the *Critique of Practical Reason* (1788) – stark oppositions are set up between duty and inclination, between morality and prudence, and between virtue and happiness. Noting these contrasts, many readers, from

¹ Kant maintains that such accounts also suffer from a more fundamental deficiency in that they fail to accommodate a conception of the will as autonomous, an inadequacy that heteronomous forms of ancient eudaimonism share with other doctrines as well, such as the utilitarian principle of universal happiness.

Kant's day down to the present, have regarded his ethics as an austere moralism, obsessed with duty and framed in a formalistic doctrine that reflects a dualistic view of human practical life. Not surprisingly, such readers have commonly found ancient eudaimonism to represent a more appealing approach.

Recent scholarship on Kant's moral philosophy has done much to counter this initial impression. It has been noted, for instance, that Kant includes a place for happiness in the highest good, and that in the *Religion within the Boundaries of Bare Reason* (1793), the *Metaphysics of Morals* (1797), and other later writings he represents that good, not as an otherworldly afterlife (as it has seemed to many that he did in the second *Critique*), but as an end to be pursued here, in the present world. In those writings we also find him acknowledging that natural inclinations considered in themselves are good, and that the virtuous frame of mind is characteristically cheerful rather than downcast or sullen.

Yet these late works were composed in the knowledge of criticisms that contemporaries had directed against the doctrines set out in the *Groundwork* and the second *Critique*.² Hence the initial impression may linger. The more accommodating attitude toward happiness may seem to be an afterthought, a concession wrung from him by the unfavorable reception of earlier, unguarded statements that provide a more revealing window on the true spirit of his philosophy.

Although I believe that a careful reading of Kant's earlier published writings can remove that impression, I shall not pursue such a reading here. The aim will rather be to show, from a consideration of student records of Kant's lectures on moral philosophy (primarily the notes of Georg Ludwig Collins³), that Kant's acknowledgment in his later writings of the value of happiness and of the importance of its inclusion in a virtuous life expresses in its essentials an understanding of morality and happiness that was firmly in place well before the publication of the *Groundwork*. It will emerge that in those early years, and in the *Groundwork*, too, Kant not

² See in particular Kant's replies to Christian Garve in *On the Common Saying* (TP 8:278–289) and to Friedrich Schiller in the *Religion* (RGV 6:23f n.). A recent discussion of the exchange with Schiller can be found in Beiser 2005, 169–190.

³ Collins attended the University of Königsberg in 1784, around the time of the publication of the *Groundwork*, but his notes are very similar to other sets of notes, which are thought to date from the mid-to-late 1770s. Included among these are the notes published in Menzer 1924, which have long been available in English translation (Infield 1930), and the Kaehler notes, recently published in Stark 2004, which deserve special mention on account of their completeness and reliability. When I speak simply of Kant's "lectures," I normally mean those recorded in the notes of Collins. My translations usually follow the Cambridge edition.

only took the highest good to include happiness as an integral component, but also followed the ancient Greeks in making the question of what the highest good consists in the starting point of his moral philosophy, and even saw his own answer to that question as in agreement with the general answer given by the ancients themselves. Far from regarding morality and happiness as fundamentally opposed, Kant held even in those early years that in the highest good the good will and virtue are internally related to happiness, constituting not only the sole condition under which happiness is good, but also the sole cause through which the good of happiness can be attained.

I

In view of Kant's reputation as a chief critic of teleological approaches in ethics and of ancient eudaimonism in particular, it is remarkable that, after a short introductory "Proem," the lectures recorded in the Collins notes enter directly into an extended discussion of the primary question of ancient ethics, concerning the makeup of the *summum bonum*, a discussion in which various positions on the question are classified and assessed. To appreciate the significance of this discussion, it will be helpful to consider its context and to draw some comparisons with the *Groundwork*.

Kant's lecture course divides into two main parts. The first is devoted to "universal practical philosophy," which concerns general topics pertaining to both ethics and the doctrine of right (e.g., obligation, law, the supreme principle of morality), roughly the territory later covered in § IV⁴ of the Introduction to the *Metaphysics of Morals* (MS 6:221–228). The second part is devoted specifically to ethics and corresponds, again roughly, to the second part of the *Metaphysics of Morals*, the Doctrine of Virtue. In both parts, the sequence of topics reflects, approximately, the order in Baumgarten's texts.⁵ At the beginning of the lectures, however, we find three sections – "Proem," "The moral systems of the ancients," and "Of the principle of morality" – which appear to serve as an introduction to the lectures as a whole and in

⁴ Or § III in the reordering proposed by Bernd Ludwig and followed in the Cambridge edition (see Ludwig 1986, "Einleitung").

⁵ Textbooks were required by the Prussian authorities, and the two that Kant used – Alexander Baumgarten's *Initia philosophiae practicae primae* (1760) and *Ethica philosophica* (1751, 1763) – expound the rational perfectionist doctrines of Christian Wolff. Kant criticizes Baumgarten on many points, but in the main follows, though not always happily, his order of presentation.

which the discussion is ordered according to an independently conceived plan.⁶ Moreover, these sections contain a sequence of thought that anticipates the organization of the *Groundwork*.⁷

Consider first the *Groundwork*. After characterizing moral philosophy as a science concerned with freedom and the laws governing what ought to happen, Kant stresses in the Preface that, as rational knowledge, it must rest on its pure, rational part, which abstracts from all special objects to which our freedom might be directed and deals merely with the universal laws of freedom that hold for all rational beings. Moral philosophy's first order of business, therefore, is to seek out the supreme principle of morality, a task to which the first two sections of the *Groundwork* are devoted. The first begins with an argument that the only thing having unconditional goodness is a good will and then develops the concept of such a will through an analysis of the more determinate concept of duty in order to arrive at a philosophical representation of morality's principle. The second section elaborates reason's entire practical application through a systematic exposition of its imperatives and by this route reaches again the same principle, which is presented in three different formulations and is said to lie in the will's autonomy; the section concludes with a classification and critical discussion of the different possible heteronomous accounts philosophers have offered of morality's principle.

Turning now to the opening three sections of the lectures, we find in the first and the third many anticipations of the *Groundwork*'s argument. Like the *Groundwork*'s Preface, the "Proem" begins by describing moral philosophy as differing from other parts of philosophy and from the rest of

⁶ Aside from the opening discussion of the distinctions between theoretical and practical philosophy and theoretical and practical sciences (C 27:243f), which can be seen as implicitly engaging the conceptions of practical philosophy and universal practical philosophy set out in the opening sections of Baumgarten's *Initia*, it is not until the fourth section – "Of active and passive obligation" (C 27:260ff) – that the lectures turn to the series of topics addressed in Baumgarten's text.

⁷ Because the *Groundwork*, as its title indicates, is intended to "lay the ground" for the *Metaphysics of Morals*, these two works can be regarded as together making up the whole of Kant's system of moral philosophy. So to the extent that the lectures' first three sections correspond to the *Groundwork*, we can think of the whole of the lectures on moral philosophy as corresponding to this complete system, though with some notable differences, prominent among which are the following: (i) The lectures' first three sections are not marked by the "subtlety" that characterizes much of the argumentation of the *Groundwork* (they do not, for instance, investigate how imperatives are possible, as does the *Groundwork* in its second and third sections) and so, not surprisingly, are not separated off from the rest of the lectures in the way the *Groundwork*'s project is distinguished from that of a *Metaphysics of Morals* (cf. G 4:391f). (ii) The lectures contain nothing corresponding to the first part of the *Metaphysics of Morals*, the Doctrine of Right. (iii) The extended treatment of natural religion at the beginning of the part of the lectures dealing with ethics (C 27:305–335) is vastly reduced in the *Metaphysics of Morals* and is moved – in keeping with the judgment expressed in the lectures themselves (C 27:305) – to the conclusion of the Doctrine of Virtue (MS 6:486–488).

practical knowledge in that it concerns the use of freedom in general, in abstraction from the specific objects we might employ our freedom to pursue. And the third section – “Of the principle of morality” – begins by announcing the intention “to see what the principle of morality consists in” (C 27:252) – that is, to do what Kant undertakes to do in the first two sections of the *Groundwork*. In content it resembles the *Groundwork*'s second section, in that it searches for the principle by way of a systematic account of imperatives of reason and includes a classification and critical examination of the systems of morality that philosophers have developed, though with the difference that the discussion of the systems comes before rather than after the account of imperatives and reveals that the principle sought after must be intellectual rather than empirical and internal rather than external.

At first glance, the lectures' second section – “The moral systems of the ancients” – may seem to lack such parallels. Its position in the lectures would appear to correspond to the location of the famous argument concerning the good will in the *Groundwork*. But what we find is an extended discussion of the ancients' ideal of the *summum bonum* and their disputes about its makeup.

On closer scrutiny, however, this difference proves to be superficial. Although the *summum bonum* is not explicitly mentioned in the *Groundwork*'s initial argument concerning the good will (G 4:393f), it comes to the fore in the argument that directly follows (G 4:395f), in which the good will and happiness are compared in order to determine which is better suited to be the primary aim that nature in its wisdom pursues in assigning reason as the governor of our will. Upon reaching his conclusion that of the two the good will is the better suited to be this purpose, Kant immediately infers that the good will “must be the highest good and the condition of all the rest, even of all demands for happiness,” even if happiness must be added to yield “the whole good” (G 4:396).⁸ When we take this second argument into consideration, then, the parallels come into view. Just as in the *Groundwork* the seeking out of the principle of morality through an analysis of the concept of duty (G 4:397ff) is preceded by an argument that locates “the highest good” by identifying the good whose goodness is unconditioned and as such the condition of the goodness of all other goods (G 4:393–396), so in the lectures the third section's investigation of morality's principle is preceded by the second section's identification of “the ideal of the

⁸ Contrary to his usual practice, Kant here uses “the highest good” (*das höchste Gut*) to signify the *supreme* good rather than the *whole* (or *complete*) good (see KpV 5:110).

greatest moral perfection” (C 27:252) through an examination of ancient accounts of the *summum bonum*.⁹ Just as in the *Groundwork* the highest good is found to lie in the good will, so in the lectures the ancients are credited with the insight that “moral perfection” lies in “the constitution and perfection of the free power of choice” (C 27:248). And just as the lectures infer from this insight that “the principle of morality . . . rests upon the goodness of the free power of choice” (C 27:252), so the *Groundwork* seeks to bring morality’s principle to light by developing the concept of a good will (G 4:397). So it turns out – given that the lectures move easily between talk of the power of choice (*Willkür*) and talk of the will (*Wille*) – that in the lectures, as in the *Groundwork*, the search for the principle of morality begins with a consideration of the highest good, which leads to the recognition that this principle is the principle figuring in the exercise of the good will.

2

It will prove instructive here to examine more closely the lectures’ discussion of the highest good. Although the section on the moral systems of the ancients is devoted largely to a critical assessment of the specific accounts of the *summum bonum* that divide the ancient schools, it begins with a generic characterization of the highest good, which is based on two insights attributed to the ancients generally. The first is that “happiness alone could not be the one highest good, for if all men were to hit upon this happiness, without distinction of the just and the unjust, then there would indeed be happiness, but no worthiness of it, and if this is taken together, that is the highest good” (C 27:247). This acknowledgment that the ancients recognized that the *summum bonum* includes not only happiness but also the worthiness of happiness and the justice on which it depends is noteworthy in view of Kant’s reputation as a critic of ancient eudaimonism. And while the suggestion that the ancients regarded the virtue of justice as rendering its possessor worthy of happiness may seem questionable (particularly if it is assumed that the notion of worthiness here introduced is the one that figures in Kant’s own account of the highest good, as usually interpreted), there is considerable plausibility in the suggestion that they generally recognized the *summum bonum* to require justice and virtue.¹⁰

⁹ Although the good will is not explicitly characterized in terms of perfection in the first section of the *Groundwork*, it is so described in the second: G 4:443; cf. 408f.

¹⁰ Although justice is the only virtue explicitly noted, there does not appear to be any intention to suggest that the ancients’ insight relates to justice alone, to the exclusion of the other virtues; in the lines that follow, worthiness of happiness is associated with “good conduct” (*Wohllverhalten*) overall,

But in the next sentence we find an even more striking proposition: "Man can only hope to be happy insofar as he makes himself worthy of it, for that is the condition of happiness required by reason itself." Given its location, this statement would appear to be a continuation or completion of the articulation of the first insight. Its precise meaning is obscured somewhat by the fact that it does not specify the sense in which worthiness is supposed to be the condition of happiness. It seems clear, however, that the intended conditionality cannot be captured in the *theoretical* proposition that happiness is simply impossible except where worthiness is present; for in the preceding sentence, as we have seen, the possibility is hypothetically entertained that there might be happiness without worthiness. Evidently, then, the condition spoken of is being conceived as a requirement of *practical* reason. The suggestion would accordingly be that reason sets worthiness of happiness as the condition of the *goodness* of happiness, and that so far as reason is efficacious, the happiness thought in its representation of this conditionality relation will be achievable just to the extent that worthiness is present. So understood, the proposition would assert that so far as man hopes in accordance with reason (so that the objects of his hope are indeed goods, as, in hoping for them, he supposes them to be), he can hope to be happy just insofar as he makes himself worthy of it. If this interpretation provides the proper specification of how worthiness, according to reason's requirement, is the condition of happiness, then we find in this statement a strengthened formulation of the initially articulated insight. Worthiness is now represented not merely as something that needs to be added to happiness to yield the highest good, but as something that must be present if reason's condition on the goodness of happiness, and hence on the eligibility of happiness for inclusion in the highest good, is to be satisfied. It seems, then, that, according to this first ancient insight, the highest good is not to be conceived as an aggregate of worthiness and happiness, but rather as a whole in which the former is, in the sense just indicated, the condition of the latter.

The ancients are also credited with the further insight that "happiness rests on the goodness of the free will, on the dispositions [*Gesinnungen*] to make use of everything that nature richly bestows on one. Of a man who is rich and has treasures in plenty, the question is, how is he disposed to make

not specifically with justice (C 27:248). That the ancients generally recognized that happiness alone, without virtue, could not be the highest good seems also to be the view of Aristotle, who says, "no one would maintain that he is blessed [*makarion*] who has not in him a particle of courage or temperance or justice or practical wisdom" (*Politics* 1323a27–29 [Barnes 1984, 2100]). How "worthiness of happiness" is to be understood will be considered in [Section 4](#) of this chapter.

use of them [*was er vor Gesinnungen hat, von diesen für einen Gebrauch zu machen*]? So the constitution and perfection of the free power of choice, which contains the ground of the worthiness of happiness, is moral perfection" (C 27:247f). This second insight deepens the first by locating the ground of worthiness in how the will is disposed to make use of things. Thus, the ancients are held to recognize not only that

- (1.1) the highest good includes not only happiness but also the worthiness of happiness, so much so that (1.2) worthiness is the condition under which happiness qualifies as good and so as an element in the highest good,

but also that

- (2.1) worthiness is grounded in those dispositions to use nature's gifts that constitute the perfection of the free power of choice, that is, moral perfection, so that (2.2) the dependence of happiness for its goodness on worthiness is at bottom a dependence on the goodness of the free will, or on moral perfection.

In the lectures' brief exposition of these two insights we find a compact expression of the same basic constellation of ideas that drives the argument concerning the good will in the *Groundwork*, notwithstanding the difference that here but not in the *Groundwork* the ideas are articulated through consideration of the ancients' insights into the constitution of the *summum bonum*. In the *Groundwork*, Kant argues that the good will is the only thing unconditionally good, being itself the sole factor whose presence or absence determines whether or not the other candidate goods – gifts of nature and gifts of fortune, including even happiness – are indeed good; and the grounds on which this distinctive status – namely of being the unconditioned condition in respect of the good – is accorded to the good will are that the will is the capacity to make use of those other things, and that a will is a good will just insofar as it is disposed, on account of its specific constitution, or character, to use them well, that is, in the way that constitutes the condition under which they are good. Here in the lectures the same point is being made about the will and its goodness. Although the argument in the lectures does not explicitly claim that the goodness of the treasures nature has bestowed on the rich man depends on the goodness of his will, it does implicitly claim (if the foregoing interpretation of the first insight is correct) that the goodness of his happiness depends on the goodness of his will, and it explicitly relies on the idea that his use of those treasures is a matter of his exercise of his free will, and on the further idea that insofar as he is disposed to use those treasures well, his will is good.

Although this comparison should suffice to put beyond doubt the similarity just described between the lectures' exposition of the ancients' insights into the constitution of the *summum bonum* and the *Groundwork's* argument concerning the good will, a further confirmation is provided when we enlarge the comparison to include the Mrongovius notes of lectures Kant offered in the months immediately following his delivery of the manuscript of his *Groundwork* to the printer. These records show clearly that the *Groundwork's* argument that the good will is the sole unconditioned good and condition of all other goods focuses directly on the question on which "the ancient Greeks" are said to have concentrated in determining the principle of morality – the question: "What is the highest good?" (M II 29:599). When the answer to this question provided in the Mrongovius notes is compared with the *Groundwork's* argument concerning the good will and with the passages from the Collins notes, it is readily apparent that Kant takes himself to be in agreement with the ancients on each of the following points:

- (i) The good will is the only unconditioned good and as such the one *supreme* good (or, as expressed in Mrongovius: "nothing is good without restriction, save the good will" [M II 29:599]).
- (ii) The *whole* good includes happiness in addition to the good will (Mrongovius: "the possession of virtue is not yet the whole good . . . virtue combined with happiness is the highest good" [M II 29:599f]).
- (iii) Happiness nevertheless depends for its goodness on the worthiness of happiness, which is grounded in the disposition of the will (i.e. in how it is disposed to make use of the gifts at its disposal) – the disposition, namely, that constitutes the will's goodness or moral perfection (Mrongovius: "well-being and constant cheerfulness of heart are good only under the condition that man has a good will, in order to make use even of them" [M II 29:599]).

From these comparisons, then, we can see that Kant regards his view of morality, happiness, and their relation in the highest good as much closer to the ancients' than is commonly supposed.

3

It should not be forgotten, of course, that in the Collins notes the exposition of the ancients' insights into the ideal of the *summum bonum* is followed by a critical discussion of their competing accounts. Three versions of the ideal are said to have been developed. The Cynics identify it with simplicity ("the man of nature"), the Epicureans with prudence ("the man of the world"),

and the Stoics with wisdom (“the sage”). Each of these accounts reflects a different determination of the way in which the ideal is attainable. Contending that the cultivation of arts and skills tends to increase our needs and so also the difficulty of attaining happiness and the likelihood of lapsing into vice, the Cynics take the highest good to lie in a simple life in accordance with nature. In opposition to the Cynics, the Epicureans and Stoics agree that this ideal can be attained only through art, but they adopt opposing views of how its components are united. The Epicureans take happiness to be the end, and noting that happiness depends on security from all reproaches from oneself or others, they represent virtue as the means. The Stoics take the wisdom of virtue as the true end and maintain that consciousness of one’s virtue is sufficient for happiness.

At the end of the section, a general criticism is raised against the three sects. It is not the charge of eudaimonism, however, or that they reduce morality to prudence or refined self-love, but that they fail to frame their ideal of moral perfection – their representation of the good will and its disposition – in its purity. Their candidate ideals, it is alleged, are not freed from the deficiency and weakness of human nature and hence fall short of the *greatest* moral perfection, which is identified as the ideal of holiness (C 27:251f, 294). The criticism is not spelled out in detail, but the problem, at root, appears to be that the ancient ideals are framed subject to the constraint that they be naturally attainable; they thus involve a tacit comparison with an empirically specified conception of human nature, which as such inevitably represents that nature as deficient. The ancient philosophers “demanded no more of man than what his nature could achieve; hence their law had no purity” (C 27:294).¹¹ And where, as in the case of the Stoics, their ideal was exalted as exceptional, as elevated “above the capacity of human nature” (that is, above what this capacity could attain without the art of Stoic virtue), the sublime conduct of such an individual was motivated by “pride, honor, etc., for example in extraordinary bravery, magnanimity” (C 27:294).¹² The ideal of holiness, in contrast, is independent of any representation of the deficiency of human nature. While this independence

¹¹ The quoted passage has been emended in the light of the editions of Menzer and Stark; in Collins, it reads as follows: “demanded no more of men than they could achieve; but their law had no purity.”

¹² In the *Critique of Practical Reason*, Kant notes that the Stoics went so far as to liken the sage to a deity (KpV 5:127; cf. C 27:250). Describing the Stoic ideal as “a certain heroism of the sage,” Kant adds that what motivates the Stoic’s sage is not the pure moral law itself, but the sage’s consciousness of the sublime elevation of his way of thinking over the lowly springs of the senses (which have strength only through weakness in the soul), an elevation on account of which the sage is fancied to be not subject to temptation (KpV 5:127n.). As Kant observes, the sage’s virtuous conduct could never be traced to such a motivational spring were the Stoic’s rule of conduct not tainted by deficiencies of human nature.

entails that holiness and the happiness or blessedness that would be suited to it are not, except through divine assistance, within the reach of beings whose nature is deficient, it also entails that this ideal is not subject to the general criticism of ancient Greek philosophy. This ideal is attributed to another ancient sect, Christianity.

Curiously, however, at the moment of transition from the discussion of the Stoic ideal to the first consideration of the ideal of holiness, yet another ancient ideal is mentioned, though only fleetingly and in passing. This other ideal is associated with Plato and characterized as “man seeing himself in communion with the highest being” (C 27:250). In the Mrongovius notes, Plato’s account is represented as different from the other three Greek ideals in that it traces the attainment of morality not to a cause in our nature but to a “supernatural” cause: “the immediate intuition of God in his ideas” (M II 29:603).¹³ Not surprisingly, the Platonic ideal is not subjected to the general criticism of impurity leveled against the ideals of the other three sects. It is criticized rather as mystical and fantastic:¹⁴ though “indeed perfect,” it is “not at all suited to human nature” and “cannot be attained” (C 27:250, 304f).¹⁵ The ideal of holiness is thus situated between ideals of two opposing types: ideals that fail to exclude *deficiencies* of human nature and an ideal that fails to accommodate the *limits* of human nature.

Kant seems prepared, however, to acknowledge the possibility of a Platonic ideal that, like the ideal of holiness, is not subject to either of the two criticisms just noted. In the *Critique of Pure Reason*, he says the exalted language Plato employs when speaking of the ideas is quite capable of a milder interpretation (KrV A 314/B 371n.). And he dismisses the commonplace objection to Plato’s ideal republic that it is visionary and impracticable,

¹³ Allen Wood suggests that what Kant represents as the Platonic ideal may be closer to the later Academy or even neo-Platonism (Wood 2005); he cites a passage from the lecture notes of Powalski (pre-*Groundwork*), where it is stated that “Plato derived all sources of our understanding from God, and said that all our happiness would consist in raising ourselves above the sensible and uniting ourselves with our source of all in the highest being” (P 27:105). Another possibility, suggested by the context of this passage, is that Kant sees Plato’s mysticism as reflecting a Pythagorean influence. In the Vigilantius notes (1793) – which classify four Greek sects under just two heads, assimilating the Cynics to the Epicureans and the Platonists to the Stoics (V 27:482–485) – mysticism is still attributed to Plato (V 27:484), even though Plato’s views are there distinguished from those of the Platonists and the neo-Platonists (cf. V 27:634f, 704, 711).

¹⁴ Or fanatical (Kachler).

¹⁵ The ideal of holiness too is said to be one that “cannot be attained by man” (C 27:251), but the thought here seems to be only that it cannot be attained by humans *alone*: it can be attained, but only with divine assistance (C 27:252, 294); it is also spoken of as an ideal that can be approached (C 27:250). Elsewhere Kant denies that the purity of morality’s ideas entails that they are anything “excessive [*Überschwingliches*], that is, something of which we could not even determine the concept sufficiently or of which it is uncertain whether an object corresponds to it at all” (KpV 5:127n.).

advising that we would “do better to follow up this thought, and (where the excellent man leaves us without help) to bring it, through new effort, to light, rather than to set it aside as useless on the very wretched and harmful pretext of impracticability” (KrV A 316/B 372f).¹⁶

Moreover, there are clear signs that Kant himself takes this advice to heart in developing his moral philosophy.¹⁷ One such indication is that, in criticizing the ancients’ naturalistically conceived ideals on account of their impurity, the lectures rely on an evidently Platonic conception of an ideal.¹⁸ According to the lectures, an ideal is “a maximum of what can be conceived of a thing, by which one determines and measures everything” and “by which everything can be judged,” so the *summum bonum*, as an ideal, is “scarcely possible” but is “a pattern, an idea, an archetype for all our concepts of the good” (C 27:247; cf. KrV A 314–317/B 371–374). The moral law is similarly characterized and even likened to geometry’s postulates in respect of exactitude, precision, rigor, and indifference to whether humans can observe the rule in practice (C 27:294, 301; cf. KpV 5:31). As the original standard for judgment, the ideal must be conceived independently of the things to be judged, which, being always to some extent deficient, never fully attain it.

In addition, an ancient version of the *Groundwork*’s opening argument concerning the unconditioned good is to be found in Platonic dialogues with which Kant seems to have been familiar. In Plato’s version, which is identical with Kant’s in basic structure, it is argued that practical wisdom is the only thing good by itself alone (or, in Kant’s terms, unconditionally good) on the grounds that a survey of all other things we might call good –

¹⁶ In the paragraph following, the tone is even more emphatic: “If one sets aside the exaggeration in expression, the philosopher’s spiritual flight, ascending from the ectypal contemplation of the physical in the world order to the architectonic connection of the same according to ends, that is, according to ideas, is an endeavor that deserves to be respected and followed, but with regard to what concerns the principles of morality, legislation, and religion . . . [it is] a quite distinctive service” (KrV A 318/B 375).

¹⁷ I say “developing” with an eye both to Kant’s statement in the first section of the *Groundwork* that he is taking up the concept of duty in order to *develop* (*entwickeln*) the concept of a good will, which “always stands uppermost in estimating the complete worth of our actions and constitutes the condition of everything else” (G 4:397), and to the following statement from Kant’s anthropology lectures: “Plato says that the preeminent business of the philosopher is to develop [*entwickeln*] the idea. This capacity to lay out [*entwerfen*] something according to an idea is reason” (VA 25:551). In the Jäsche *Logic* § 37n1 (L 9:111), *Entwicklung* is glossed as *explicatio*: making what is implicit explicit, as in a nontautological analytic judgment.

¹⁸ As Klaus Reich notes, Kant’s appreciation of the purity of the Platonic ideal is in evidence already in 1770, in § 9 of his Inaugural Dissertation, where he says that “the *maximum of perfection*, which is now called an ideal, was called by Plato an idea (as in his idea of the state)” (AA 2:396). Reich sees the publication of Mendelssohn’s *Phädon* (1767) as playing an important role in awakening Kant’s interest in Plato; see Reich 1939.

whether they lie outside the soul (e.g., wealth, health) or are qualities within it (e.g., boldness, mental quickness) – reveals that all such items are good only to the extent that they are properly used or directed by wisdom: their goodness, in short, depends on correct use, which depends in turn on wisdom.¹⁹

At first glance, Plato's version may seem to differ from Kant's in a significant respect, in that it locates the unconditioned good not in the good will, but in practical wisdom and knowledge.²⁰ Kant makes amply clear, however, that he conceives of the will as a cognitive capacity. In the *Groundwork*, he takes up the good will as part of his investigation of "moral rational knowledge" (G 4:393, cf. 387f), explicitly identifies the will with practical reason (G 4:412; cf. MS 6:213), and notes that practical reason is just reason itself, the highest cognitive power, in its practical application (G 4:391; cf. KpV 5:121). And in the lectures, it is stated that practical philosophy is "the science of the rule of how man ought to conduct himself" (C 27:244), and that the principle of morality – the first principle of that science – "rests on the goodness of the free power of choice" and "flows from the ground of our will" (C 27:252).

So rather than opposing Plato's argument, the *Groundwork's* version highlights the practicality of the knowledge under consideration. Because the goodness of all other things depends on their correct use, the standard or principle by which their goodness is measured and determined must inform the disposition of a good will, for it is in the exercise of such a will that the correct use of them consists. Thus, by beginning his moral philosophy with this Platonic argument, Kant shows the principle of morality to be the principle of reason's practical-cognitive use, or, what comes to the same, the principle operative in the good will's exercise.

In sum, then, it appears that, although the discussion of ancient systems in the Collins notes concludes with a general criticism of the Cynic, Epicurean, and Stoic sects, the criticism is so far from touching Plato that Plato has rather some claim to be a source of inspiration for it. Nor does the general criticism touch the generic conception of the *summum bonum* – comprising points (i)–(iii) – that was expounded through the two ancient insights at the outset of that discussion. Indeed, in both the lectures and the

¹⁹ See, for example, *Meno* 87d–89a (Cooper 1997, 887f); cf. *Euthydemus* 279a–281e (Cooper 1997, 716–719); *Republic* 504a–506a (Cooper 1997, 1125–1127); *Laws* 661a–c (Cooper 1997, 1352).

²⁰ I say "significant" in part because this difference might seem to reflect a rejection on Kant's part of the classical idea that morality is based in knowledge of the good. Readings of Kant that portray him as rejecting that idea can be found in MacIntyre 1984, 154f, and Schneewind 1998, 508–530. I compare Kant's and Plato's arguments more closely and discuss this apparent difference in Engstrom 1997.

Groundwork, the setting out of this conception serves as the starting point in the search for morality's principle. And since, as we have now seen, this generic conception finds expression in Plato's dialogues as well, we can also see how in his practical philosophy Kant carries out his recommendation in the first *Critique* to "follow up" Plato's idea and "to bring it, through new effort, to light."

4

In addition to countering the common depiction of Kant's moral philosophy as radically opposed to ancient ethics, the appreciation we have now gained of how ancient insights inform his account of the *summum bonum* – as comprising worthiness and happiness, related as condition and conditioned – positions us to ward off certain misconceptions of that account. In particular, we can see that, since those insights originally reside in the *practical* use of reason, Kant's account cannot be understood through the lens of the commonly held assumption that it arises within a *retrospective* view of practical life (a questionable assumption in any case, which would allow into Kant's ideal the very impurity he rejects). But because this assumption is fueled in large part by a related misconception pertaining to his idea of worthiness of happiness, dislodging it will require the removal of that misconception as well.

Kant's idea of worthiness figures prominently in the parts of his published writings that concern the highest good's relation to the postulates of pure practical reason, and in particular to the postulate of the existence of God, so it is not surprising that readers are prone to interpret his talk of worthiness in a way that directly connects the idea with a doctrine of moral desert. Such an interpretation is bolstered by Kant's description, in such contexts, of the relation between the components of the highest good in terms of a distribution of happiness in proportion to morality or virtue. Although this description expresses the relation he intends, it can elicit a distracting assemblage of ideas associated with the notion of a final judgment, as represented, say, in Christian eschatology. Even in the second *Critique*, such ideas form no part of Kant's exposition of the concept of the highest good, which situates the concept in practical reason's basic, wholly prospective use, where it represents "the ultimate object of all conduct" (KpV 5:129). But readers distracted by them may confuse that concept with a different notion that might be derived from it – the idea of a rule of desert that would govern a retrospective dispensation of rewards and punishments.

An advantage gained from noting the affinity between the *Groundwork's* argument concerning the good will and the lectures' corresponding discussion of the ancients' conception of the *summum bonum* is that we are better able to appreciate that this conception is thoroughly practical and as such wholly prospective, wholly independent of retrospective considerations. Such appreciation positions us to find a practical interpretation of Kant's talk of worthiness of happiness, one that rules out the interpretation of it in eschatological terms.

The practical interpretation can be located by considering more closely the claim in the lectures that the good will's *disposition to use* the treasures bestowed on one by nature, which is said to constitute moral perfection, "contains the ground of the worthiness of happiness" (C 27:248: see [2.1] in [Section 2](#) of this chapter). This claim is not spelled out, but the comparison of the good will and happiness in the *Groundwork's* opening paragraph furnishes some clues. There two points are noted. The second, that "a rational and impartial spectator" cannot delight in the sight of the happiness of a being with no trace of a good will, may, like Kant's talk of proportionality in his account of the highest good, seem to suggest a final judgment. But this second point ultimately depends on the more fundamental first point, that happiness and the other gifts of fortune "generate boldness and thereby often overboldness where there is not present a good will, which corrects and makes universally purposive their influence on the mind and herewith the whole principle of action" (G 4:393). This point is intended to show that happiness is not unconditionally good. But it also indicates that the dependence of the goodness of happiness on the good will is grounded in the way the good will *uses* that happiness, or in the way it manages its good fortune.²¹ Kant's first point thereby positions us to appreciate a familiar notion of worthiness, one that the following contrast can bring into clearer view.

When we speak of a person as worthy of a certain thing, one thing we can mean is that they deserve it in recognition of, and perhaps as a reward for, something they have done or the way they have conducted themselves. This sense is normally in play when we describe someone as worthy of praise. Such worthiness entails no relation of use between the action or conduct and the thing. But another thing we can mean is that the person is disposed to use or to treat the thing properly. This is the sense we intend when, for

²¹ Essentially the same point is made in a passage quoted earlier from the Mrongovius notes: "well-being and constant cheerfulness of heart are good only under the condition that man has a good will, in order to make use even of them" (M II 29:599).

instance, we speak of a person as trustworthy (*vertrauenswürdig*). To be worthy of trust is to be disposed to treat properly the trust others place in one. Or again, to be worthy of a certain office is to be disposed to carry out the responsibilities attached to it. Similarly, we can say that worthiness of happiness lies in the disposition to manage properly such prosperity as one may enjoy, including all the means at one's disposal in securing and maintaining it. We find this sense of 'worthy' in play in one of Kant's explanations of what worthiness to be happy consists in, where he says that if the will of a person who has "skill in achieving some happiness" conflicts with morality, then he "is not worthy even of this skill and the talent that nature has lent him for it" (TP 8:278n.).

Worthiness in this practical sense lies not in capacity but in disposition. It depends not on the extent of ability, but on the quality of will. Both of these factors must of course be present if a person is to be fully effective. But worthiness relates to will and its quality, not to capacity and its extent. This difference is explicitly marked later in the lectures, when a distinction is drawn between perfection of one's *capacities* and (*moral*) *goodness*, or perfection of the *will*. Perfection of capacity "can be greater or smaller . . . But goodness is the property of making good and proper use of all these perfections. Moral goodness therefore consists in the perfection of the will and not of capacity."²² Worthiness, then, consists in a quality of will, namely the will's being disposed to make good and proper use of all the materials at its disposal. In the *Groundwork*, Kant calls these materials "gifts" (*Gaben*) of nature and of fortune (G 4:393),²³ suggesting thereby that they have been *entrusted* to the will and further that they have the *potential*, on the condition that they are properly used, to be good, or, what comes to the same, to contribute to the *summum bonum*.

Once we have distinguished this practical sense of 'worthiness' from the sense the term can have in retrospective assessment of conduct, we can appreciate that insofar as it is worthiness in this sense that is understood to be the condition of happiness in the *summum bonum*, this ideal is being conceived not as a final outcome that follows as a *consequence* of an assessment of conduct but as an end that, prior to any such assessment, lies internal to practice. It is an end in which, through worthiness, or the goodness of one's willing, one does not allow such happiness as one enjoys

²² C 27:265f (cf. MS 6:391–393). Because Kant's account of the good will is often faulted for overlooking the good will's crucial reliance on competence in order to achieve its ends, it is worth noting that this obvious point is acknowledged in the immediately following sentence: "Yet a good will needs the completeness and capacity of all powers to carry out everything the will wills" (C 27:266).

²³ The lectures similarly speak of what nature "bestows" (*schenkt*) on one (C 27:247).

to influence one to engage in conduct not in accordance with morality's principle – "the rule of how man ought to conduct himself" (C 27:244). So conceived, the highest good is a whole of human activity in which the conditioned good does not influence the disposition of the will that constitutes the condition of that conditioned good's goodness. The causal relation that Kant asserts between virtue and happiness (KpV 5:111) is here regarded in such a way that happiness is understood, not just as a consequence of virtue that results indirectly through the administration of divine justice, but as the conditioned, material element in the *summum bonum* conceived as a practical whole, in which the conditioning element serves a direct regulating function, through use.

This pivotal role that the notion of use plays in Kant's conception of the *summum bonum* brings into view that the conditioning relation the unconditioned good bears to the conditioned good is one of internal, or formal, causal determination. The good will is not only the condition of the goodness of happiness, but also the only possible sufficient cause of happiness itself so far as the latter is anything good. Thus, as we are now in a position to see, certain striking statements that we noted in the lectures' exposition of the two ancient insights – namely, "Man can only hope to be happy insofar as he makes himself worthy of it, for that is the condition of happiness required by reason itself," and "happiness rests on the goodness of the free will" (C 27:247) – have an immanent-causal as well as rational-cognitive significance. Since humans depend on social and natural conditions, the causal relation between the two elements of the highest good can be represented as *complete* – as it is when this ideal is represented in its purity – only so far as it is represented as holding not only distributively, in the individual, but also collectively and with divine assistance (as expressed in the postulates of pure practical reason). But this relation is still conceived as one in which the will's goodness is itself the direct cause of happiness in the highest good. This idea finds clear expression in the final section of the lectures, where it is stated that "God wills not merely that we should be happy, but that we should make ourselves happy, and this is the true morality" (C 27:470; cf. 285f). Kant's moral philosophy thus ends where it begins, with the *summum bonum*, conceived as a purely practical end, one in which the good will is the condition of the possibility of happiness and also its sufficient cause.

*Kant's history of ethics**Allen W. Wood***1 Kant and the history of reason**

Kant was not a knowledgeable historian of philosophy. He came to the study of philosophy from natural science, and only later in his career did the fields of ethics, aesthetics, politics and religion come to occupy his central concerns. He never gave the history of philosophy the prominence in his works that Hegel (for instance) did, though as we shall see, this is more a matter of form and presentation than of the substance of his thought. Kant was well acquainted, of course, with the recent tradition of German philosophy: Leibniz, Wolff, Baumgarten, and Crusius, and he seems also to have had knowledge of eighteenth-century French philosophy, and of as much of Locke, Hume, and other Anglophone philosophy as had been translated into either French or German. But Kant's knowledge of classical Greek philosophy was mostly at second hand (mainly through Cicero's Latin popularizations and via J. J. Brucker's *Historia critica philosophica* [1742–1744]). Like many modern philosophers (even Hegel), he had a woefully inadequate appreciation of the scholastic philosophy of the high middle ages, even though his own philosophy, especially his philosophical terminology (via Wolff and Baumgarten), was deeply indebted to it.

Many readers of Kant, especially those under the influence of Hegel, even think that Kant's philosophy is "ahistorical." Sometimes the claim takes the form that Kant did not conceive of *reason* historically (something that Hegel, by contrast, is supposed to have done). I have long thought that this claim is completely wrongheaded – in fact, doubly mistaken. For, perhaps beginning from a crude and false conception of all these matters, it ignores or mistakes the subtle relationship between history and (timeless) truth in both thinkers. First, it exaggerates the historicism of Hegel's philosophy, whose foundation in fact lies in the "thought-determinations" of speculative logic, which Hegel conceives nonhistorically (even atemporally), as having timeless validity for both thinking and being. I do not mean

this as a reproach – and Hegel would never have regarded it as such. To him it would mean only that philosophy deals with what is in the highest sense true – with God or the absolute – rather than with what is merely transitory and contingent (see Hegel 1991, §§ 1, 6).

Second – and perhaps more to the point – to contrast Hegel with Kant in this way also ignores the ways in which Kant's philosophy *is* historical even in its self-conception – ways in which Kant actually anticipates many of the very features of Hegel's philosophy that lead people to describe Hegel as having a “historical” conception of philosophy. The common blindness to the historicity of reason in Kant is a long-standing obstacle to our understanding of Kant's philosophy.

Both Prefaces to the *Critique of Pure Reason* describe the task of that work in precisely historical terms – the A Preface by an elaborate political analogy with the ongoing modernization of European states, the B Preface by analogy with the history of science (in particular, the sciences of logic, mathematics and physics) (KrV A viii–xiii, B vii–xvi). Kant's philosophical reflections on both politics and religion rest on a historical conception of the state and the church, and are self-consciously designed for an age of enlightenment.

Moreover, Kant's conception of human nature itself – “the character of the human species” – and that means the nature of reason, as the uniquely self-creating and self-developing faculty of the mind, is historical through and through. At the end of *Anthropology from a Pragmatic Point of View*, the character of the human species is presented as “a character that [the human being] himself creates, insofar as he is capable of perfecting himself according to ends that he himself adopts. By means of this the human being, as an animal endowed with the capacity of reason (*animal rationabilis*) can make out of himself a rational animal (*animal rationale*)” (ApH 7:321). Human rationality is in this sense always something self-created and self-creating, through an endless historical process. This rationality, as Kant tells us in *Idea for a Universal History*, is developed only historically, “*only in the species and not in the individual*” (I 8:18; cf. ApH 7:329). “The first character of the human being is the capacity as a rational being to obtain a character as such for his own person and for the society in which nature has placed him” (ApH 7:329).

The development of the species-character through reason is precisely the collective historical vocation of humanity, consisting in the open-ended development of its rational species capacities. Kant divides these into the *technical* predisposition (the endless development of skills in devising means to freely self-given ends), the *pragmatic* predisposition (the restless search for happiness in an ever-changing social context), and the *moral* predisposition

(which Kant thinks has barely begun its development, far from having completed it) (ApH 7:322–324; MA 8:116f). These predispositions are themselves presented as a set of collective historical tasks for humanity, the threefold process of “cultivation,” “civilization” and “moralization” (ApH 7:325; cf. I 8:26). Human nature for Kant, which means the nature of reason itself, is nothing fixed by nature, but only the historical product of self-creation and self-education (ApH 7:325–328). All our attempts – for instance, Kant’s own attempt in the *Dialectic of Pure Reason* – to expound the principles of reason, must always be understood as part of an ongoing process of rational communication, on which the very existence of reason depends (KrV A 739/B 766).

In short: to anyone properly acquainted with Kant’s writings on anthropology and history, the common charge that Kant’s philosophy is “ahistorical” is not merely oversimplified or one-sided; it is fundamentally, wildly, and conspicuously false.

2 Kant and the history of ethics

In Kant’s ethical works in particular, however, Kant’s historical references seem at first glance only occasional, not systematic. He sometimes compares or contrasts his position with that of the Wolffians, the Stoics or the Epicureans, but his most conspicuous historical references are in his systematic account, in the *Groundwork* and the second *Critique*, of the way in which previous moral philosophers had conceived of the supreme principle of morality as a principle of heteronomy, in contrast to the principle of rational autonomy through which he proposes to reform the foundations of practical philosophy. But this reference itself looks more like a repudiation of the entire history of ethics than like a historically self-conscious way of conceiving his own ethical theory. Yet if we know where to look, we find that Kant did have an interesting and distinctive view about the history of ethics, which sheds considerable light on the relation of his own moral philosophy both to the moral philosophy of the ancient world and to other moral philosophies of his own time. It turns out that Kantian ethics is equally historical in its conception, standing in a similarly self-conscious relation to the history of philosophical ethics as Kant conceived it. We do not find this history presented explicitly or systematically in any of Kant’s published writings on ethics (in them it is adumbrated only in a single footnote, KpV 5:127n.). But when we turn to the transcriptions of Kant’s lectures on ethics, we find that throughout his career, Kant regularly began these lectures with a brief survey of the history of ethics. It was, no doubt,

presented with the intention of providing his students with a routine overview of the history of the subject matter. At the same time, however, we can also see how Kant is using his historical introduction to motivate his own original approach to the topic of searching for a supreme principle of morality. By looking at this history, we can come to understand Kant's own enterprise in ethics as a projection of certain vital historical developments in ethics, as Kant sees them. If to conceive philosophy historically is to recognize that philosophers, and human culture generally, thought quite differently about things at different times, and to see the historical development of this thinking as a progressively deepening understanding,¹ then Kant conceived of moral philosophy, and even of moral *reason*, historically, every bit as much as Hegel did. And like Hegel, he saw Christianity as playing a pivotal role in the course of historical development.

3 Kant's taxonomy of moral principles

The best place to begin, however, is not at the beginning (that is, the beginning considered in itself, the historical beginning, which Kant locates in ancient Greek ethics) but rather with what is better known to us (as readers of his published ethical writings), namely, his discussion of previous attempts to formulate the supreme principle of morality. What Kant offers us is a taxonomy, simpler in the *Groundwork*, more complex in the second *Critique*, with some interesting embroidery in some of his lecture presentations. We may summarize this taxonomy in the following table (cf. G 4:441–444; KpV 5:40; C 27:100, 253; V 27:510; M II 29:621f, 625–627).

<i>Principles of heteronomy</i>	
<i>Subjective (Empirical)</i>	
External	Internal
Education (Montaigne, [Mandeville])	Physical feeling (Epicurus, [Hélvetius, Lamettrie])
Civil Constitution (Mandeville, [Hobbes])	Moral feeling (Hutcheson, [Shaftesbury])
<i>Objective (Rational)</i>	
Internal	External
Perfection (Wolff, the Stoics, [Baumgarten, Cumberland])	The will of God (Crusius, the theological moralists, [Baumgarten])

¹ Hegel expresses things this way, for example, in Hegel 1991, § 13.

If this systematic account of previous principles of morality counts as a “history,” it does so only when “history” is used not in the narrative-chronological sense, but in the taxonomical sense (as in Aristotle’s *Historia Animalium*). As Kant uses the term “history” in the final brief chapter of the first *Critique* (“the history of pure reason,” KrV A 852–855/B 880–883), that use is poised delicately between these two senses, so that we are not sure in which sense he is using it; and the same might be said of the above taxonomy of moral principles. Of course in both cases, Kant is presenting us with a history of *failed* attempts: in the first *Critique*, at grounding human cognition, in the *Groundwork* and second *Critique*, at formulating the supreme principle of morality. All the principles listed in the latter taxonomy are principles of heteronomy, which derive morality from something other than the will of the rational being itself.

Yet what is clear from Kant’s discussion in the *Groundwork* and even more from his treatment of these principles in his lectures, is the fact that for Kant these proposed principles of morality, although none of them is adequate, form a sort of hierarchy of approximation to an adequate principle. The “subjective” or “empirical” principles are farther from being adequate than the “objective” or “rational” ones, and the “external” versions of each kind of principle are less adequate than the “internal” ones (G 4:441–443; KpV 5:41; C 27:108–110, 252–255; M II 29:621–628). Kant therefore presents us in a sense with a kind of developmental hierarchy, not unlike the later idealist transcendental progressions used to systematize philosophical materials in Fichte’s *Wissenschaftslehre*, Schelling’s systems of idealism and natural philosophy, or Hegel’s logic, philosophy of nature, and philosophy of spirit.

The “inner” principles all come closer to the principle of autonomy by displaying the moral goodness of the action as something arising from our own will. Thus an external empirical system depending on civil constitution, since it employs external or coercive motives, is less adequate than one involving education, which aims at developing inner moral grounds (M II 29:621). The empirical principles all locate the ground of morality in something external to this will – some object or incentive presented to it, for which it has a contingent, empirical inclination. None of them can account for the categorical nature of moral imperatives, whose obligation on the will cannot be undone simply by giving up or deciding not to satisfy some desire (however urgent or central the desire may in fact be to us or our well-being).

The rational principles, in fact, even admit of an interpretation on which they might be compatible with an ethics of autonomy: if we obey the divine

will not out of empirical feelings of love or fear, but because God's perfect will commands what we objectively ought to do, or if we seek the perfection precisely of our volition as rational beings, then these principles might get it right in practice about what we ought to do. But on that interpretation they "pass the buck" both about the fundamental reason why we should do it and about the principle on which we are to act. (What exactly is it that God's perfect will wills, and why is it that we are obligated to do that? What does perfection of our volition consist in, and what is it about precisely that sort of volition that makes it obligatory for us?)

The rational principles, when so understood, have an affinity with another set of proposals about the moral principle, which Kant rejects because they are analytic and therefore provide no determinate principle at all for action.

1. Do good and avoid evil. (Wolff).
 2. Act according to the truth (Cumberland).
 3. Act according to the mean between vices (Aristotle). (C 27:264, 276f).
- "Do good and avoid evil" is trivial because the concept of a good action is simply that of an action that is to be done, and the concept of an evil action is that of one that is to be omitted. The principle attributed here to Richard Cumberland is actually one that is held, in various forms, by virtually all adherents of the British rationalist tradition in ethics, including Ralph Cudworth, Samuel Clarke, John Balguy, William Wollaston, and Richard Price.² It holds that actions have a real nature, and are involved with real relations to things and to other actions. In virtue of these natures and relations, it is *true* of some actions that they are *right* or *to be done*, and of others that they are *wrong* or *to be avoided*. Presumably Kant's criticism of the principle that one should act in accordance with such *truths* is that this principle actually says no more than Wolff's principle does (for it tells us only to perform those actions of which it is *true* that they are right and ought to be performed). It is curious that Kant should have listed Aristotle's principle of the mean along with principles of the moderns, and curious also, as we shall see in a moment, that Aristotle finds no place at all in Kant's account of the ancient schools of ethics. But his criticism is no doubt that since a mean is defined as the action that is to be done located between two

² The emphasis on truth is especially identified with Wollaston, whose views were prominently criticized (though not using his name) by Hume 1967 [1739–1740], 461. It is even a bit odd that Kant should identify this view with Cumberland (even though it is stated prominently in the opening chapter of *De Legibus Naturae* [1672]), since Cumberland is more often thought of as an ethical eudaimonist. The probable explanation is that Cumberland is the only one of these authors who wrote in Latin, and Kant did not read English.

actions that should not be done, it, like Wolff's principle, actually tells us only to do those actions that fall under the concept "to be done."³

Especially noteworthy, however, is Kant's preference for the principle of moral feeling over that of physical feeling or happiness, because it captures (albeit inadequately) the recognition that practical reason produces the direct desire to do actions that accord with the law, as well as moral feelings of approval regarding such actions. Kant was always attracted by the theory of moral sense, as represented by Shaftesbury, Hutcheson, Hume and Smith, and consistently makes a place for moral feeling (as a direct influence of reason on sensibility) as part of his moral psychology (G 4:402n.; KpV 5:80, 91f; TP 8:283; MS 6:399–403). This theme in Kant's ethics should not be underestimated, because it arises from the fact that Kant regards the principle of moral feeling as a genuine *historical* advance in the history of ethics. It, he says, "stems from the feeling for what is good as such. This is an invention of the modern age" (M II 29:623). He thus credits the moral sense school with being the first one to appreciate that ethics is grounded on the inherent worth of good actions themselves rather than on their serviceability to the good ends they seek to achieve.

In other words, moral sense theory for Kant is the historical origin of the idea that what has essentially moral worth is the good will itself rather than whatever good results it may have (G 4:394). This is at the same time the recognition that *moral*ity itself – in the shape of a good will, especially the distinctively human form of the good will, which acts from duty – has a distinctive kind of value, qualitatively different from all other goods. This for Kant is fundamental to the way modern ethics differs from ancient ethics.

However, we will not fully understand Kant's view about the essential difference between ancient and modern ethics until we leave behind his taxonomy of principles and consider instead a distinction he makes between two very different approaches to ethics, one of them essentially ancient, the other essentially modern – yet arising in his view out of a development in late antiquity, namely the rise of *Christianity*.

4 Ancient ethics as ideal ethics

It is significant that Kant's historical allusions in his systematization of heteronomous moral principles are mainly to modern moral philosophers.

³ In the *Metaphysics of Morals*, Kant criticizes Aristotle's principle not on the ground that it is analytic, but on the ground that it is *false*, since (he argues) it gives a false account of what virtue is (MS 6:404).

For Kant thinks that conceiving of ethics in terms of a principle of morality is a modern way of thinking about it. "What, then, is the basis of morality? This question has been investigated in the modern age. The principle of morality, or the logical principle, is that from which all moral laws may be derived" (M II 29:62of). By contrast, the ancient view of ethics conceives the basis of morality as an ideal of the highest good. "All ethical systems of the ancient world were founded on the question of the *Summum Bonum* and what it consists of, and the systems of antiquity are distinguished according to their answers to this question. This *Summum Bonum* I call an ideal" (C 27:247). "The ancients concentrated the determining of the principle of morality on the question: 'What is the highest good?'" (M II 29:599).

It is true that Epicurus and the Stoics are mentioned regarding two of the six kinds of principle, but from Kant's lectures we learn (perhaps to our surprise) that strictly speaking no ancient ethical theory was primarily oriented toward the formulation of a *principle* of morality at all. Instead, Kant tells us, ancient ethics was founded on a set of competing conceptions of the *summum bonum* or highest good. In ancient ethics, however, the highest good does not refer to an end to be produced by moral action that involves two heterogeneous goods related conditionally: morality of character, or worthiness to be happy, combined with happiness proportionate to that worthiness. It refers instead to a single homogeneous good – a way to live well, a way to be happy or *eudaimon* – in which there is as yet no ultimate distinction between the moral *good* of one's person or one's manner of acting and the *well-being* of one's state or condition (*Zustand*) (cf. KpV 5:57–63). Ancient ethical theories were based on the highest good in the sense of the ethical "ideal" – the *ideally best kind of person*, taking it for granted that being this kind of person is at the same time the ideally best way to live well or be happy.

In other words, Kant regards ancient ethics as an ethics of *ideal being* – what might nowadays be called a "virtue ethics" – whereas modern ethics, by contrast, as an ethics of moral doing or an ethics of *principles* – an ethics grounded in principles saying *what to do* and *why to do it*. For modern ethics, the moral good lies in the good of one's person and way of acting, while the well-being or happiness consists in a desirable state or condition. And the moral good cannot be defined prior to or independently of the moral principle telling us how to choose and act (KpV 5:63–66). Kant respects both kinds of ethics, and treats the moral ideal or highest good as still on the agenda of moral philosophy, but regards the ancient kind of ethics, regarded as the *foundation* of ethics, also as unsuited to modernity.

The deeper orientation of modern ethics he regards as having arisen through progressive developments that occurred within ancient ethics itself.

There is clearly still a place for ideals in modern ethics, in Kant's view. He endorses the Christian ideal in particular, as well as the Stoic ideal of the sage, and also the Epicurean ideal of the virtuous and cheerful heart, though he regarded both these ancient ideals as one-sided in an important way (RGV 6:60; KpV 5:111–113). Kant's endorsement of the Christian ideal presents clearly his view about the relation of ideals to principles, for he calls it "the personified idea of the good principle" (RGV 6:60). An *ideal* is the concept of an individual being (here, an individual human being) that corresponds to (or personifies) an a priori concept of reason, or an *idea*. But an idea, in turn, rests on a principle of reason (here, the *good principle*, that struggles against the radical evil in human nature). It would be a topic for another paper to explain why Kant regards ideas, hence also ideals, as grounded on principles, rather than the reverse (but see KrV A 298–332/B 355–390; KpV 5:57–63). The modern ethical theory, oriented toward *principles*, hence toward what to do and why, is therefore deeper and historically superior to the ancient ethical theory of being and virtue, oriented toward ideals. However, before we inquire after the *historical* ground of this superiority, we need to look at ancient ethics itself, as Kant understands it, and the competing conceptions it offers of the ethical ideal.

Kant distinguishes five different ethical ideals in antiquity, the first three focusing on our natural powers, and the last two involving our relation to the supernatural:

- I. The *Cynic* ideal (of Diogenes and Antisthenes), which is *natural simplicity*, and happiness as the product of nature rather than of art.
 - II. The *Epicurean* ideal, which is that of the *man of the world*, and happiness as a product of art, not of nature.
 - III. The *Stoic* ideal (of Zeno), which is that of the *sage*, and happiness as identical with moral perfection or virtue.
 - IV. The *mystical* ideal (of Plato), of the visionary character, in which the highest good consists in the human being seeing himself in communion with the highest being.
 - V. The *Christian* ideal of holiness, whose pattern is Jesus Christ.
- (C 27:100–106, 247–250; V 27:483–485; M II 29:602–604).

We may see in this list a hierarchy or progression, with the later items representing higher or more adequate conceptions of the ideal. The progression goes from a happiness that is simply natural through one that is cultured to one that approaches the supernatural, and then beyond happiness, from an ideal of knowledge and communion with the supernatural to

the ideal of achieving supernatural perfection. But each item on the list also captures in its conception of happiness something Kant regards as one part or aspect of the truth, so that the different ideals also complement one another. Although he does not say so, the first three ideals would seem to correspond to the three predispositions Kant distinguishes in human nature: animality, humanity, and personality. Thus Kant's treatment of all the ancient ideals, even the Cynic one, is more favorable than critical.

The Cynic ideal is that of innocence, separation from the misery and corruption of human society, and freeing oneself from the burdens of artificial needs and inclinations. Thus Kant associates it in the modern world with Rousseau, "that subtle Diogenes" (C 27:102, 248; V 27:484; M II 29:603). The Cynics "posited the greatest good in the *abstine*, i.e. the pleasure of being able to do without, and thus the enjoyment of life under the fewest possible requirements . . . Hence their symbol was the club of Hercules, signifying strength of mind with self-sufficiency" (V 27:484). I suggest we read Kant's view that Cynicism captures part of the truth (the least adequate part, as it will turn out) as a limited endorsement of Kant's claim in the *Groundwork* that "inclinations have so little worth in themselves that it is the universal wish of every rational being to be free of them" (G 4:428, 454). Thus Kant says that the Cynic ideal "were it attainable, would be preferable on the system of Diogenes even for the Epicurean, since there is more pleasure contained in doing without than in the burden of all the means acquired for the purpose" (V 27:484). Another Cynic element in Kant's own ethics is his adoption of the Rousseauian idea that human misery and moral corruption are products of the social condition, giving rise not only to competitiveness and hostility between people but also to the destructive social passions: tyranny, greed and ambition (see Wood 1999, 259–265).

Both the Epicurean and Stoic ideals, in contrast to the Cynic ideal, view happiness as a product not of nature but of human art. Among the moderns, Kant occasionally associates the Epicurean ideal with the French materialists Lamettrie and H elvetius (C 27:100), but sometimes also with Hume (or else – the lecture manuscripts leave this a bit unclear – with Henry Home, Lord Kames) (C 27:102, 249). The former associations tend to provoke condemnations of the ideal from him ("this is, so to speak, the philosophy of rascals"), the latter more favorable associations: that virtue must be taught, and is the possession of the cultivated man of the world.

Given his usual reputation as hostile to our natural inclinations, we might think that Kant would be fundamentally hostile toward the Epicurean ideal of the life of pleasure and fundamentally favorable toward the Stoic ideal of

a life guided by reason. In fact, however, a survey of his remarks on these ancient schools both in his published writings and his lectures reveals that he often goes out of his way not only to praise Epicurus, but also in the same connection to criticize the Stoics. One of Kant's consistent themes is that Epicurus has been misunderstood. Epicurus' ideal, he says, was that of "an inner contentment and a cheerful heart. One must be secure against all reproaches from oneself or others – but that is no philosophy of pleasure, and he has been poorly understood. We still have a letter from him, in which he invites someone to dine, but promises to receive him with nothing else but a cheerful heart and a dish of polenta – a sorry meal for an epicure" (C 27:250; cf. 27:101). The true ideal of Epicurus, according to Kant, was not the life of bodily pleasure but the cheerfulness of virtue and the self-contentment of an upright character. "Pure mental enjoyment was the pleasure that arises from the performance of virtuous acts" (V 27:483). This is the true superiority of the Epicurean ideal over the Cynic ideal, since the Epicurean seeks happiness in the exercise of reason and the development of our faculties. It is an ideal of civilization or of the "man of the world," whereas the Cynic places the ideal only in the happiness of nature (V 27:484).

It is no doubt nevertheless true that Kant regards an ethical ideal that emphasizes rational virtue as closer to the truth than one that seeks happiness in pleasure (of any kind). But more fundamentally he regards the Epicurean and Stoic ideals as each capturing one side of the truth, and in a sense as complementing one another, or – to look at it the other way – as displaying opposed one-sidednesses or deficiencies. The Epicurean seeks to identify a happy state (*Zustand*) with the self-contentment that comes from the worth of our person, whereas the Stoic looks to virtue in our person to produce a happy state. Kant, however, insists that the worth of our state is something fundamentally distinct from the worth of our person. Happiness is the former, while virtue is the latter. The Epicurean and the Stoic systems thus founder together on their failure to draw the distinction.

The confusion of morality with prudence, or ethical reason with pragmatic reason, is the basic error of which both schools are guilty. It is the fundamental error of all ancient ethics, enabling it to be an ethics proposing a single ideal way of being or highest good, in which the goodness of our person is not distinguished from the well-being of our state, but both are taken up (or confused) in the idea of eudaimonia, living well, being the ideally best kind of person.

This confusion between morality and prudence is one which Kant regards as a common human foible: "The human being can be contented

or discontented with himself either pragmatically or ethically. But he very often confuses the two. He often thinks he has pangs of conscience, although he is only afraid of a tribunal of prudence" (C 27:251). This is also, in Kant's view, the typical fault of the self-tormenting repentant sinner – "a torment which is very ambiguous, and usually only an inward reproach for having offended against prudence" (RGV 6:24). Thus people still speak of "feeling guilty" when they smoke or overeat; the Kantian diagnosis of this curious tendency is that the self-deceptive burden of moral guilt – in which we may portray ourselves in the more dignified role of *transgressor* instead of having to admit that we have been merely a *fool* – is less humiliating to them than the honest admission that they have simply behaved foolishly and harmed their own interests. But this comical form of the confusion can arise historically only after the moral good has been distinguished from well-being. In ancient ethics, by contrast, it shows itself in more naïve and loftier forms that appear to us moderns as merely strange and paradoxical: the Socratic paradox that no one can knowingly do evil, or the Stoic's insistence that the virtuous person is happy even while being tortured on the rack.

5 The supernatural ideal

The Cynic, Epicurean, and Stoic ideals identify the ethical ideal with *happiness*. The supernatural ideals of Platonism and Christianity, however, resolve the contradiction within and between Epicureanism and Stoicism by breaking entirely with the eudaimonism of the natural ethical ideals. This is, to be sure, anticipated already in the Stoic ideal: "The archetypal idea of Zeno is the sage, who feels happiness within himself, who possesses everything, and who has in himself the source of cheerfulness and righteousness . . . [He] is ranked above the gods, since divinity had no temptations or obstacles to overcome; but a sage of that kind would have attained to such perfection by his strength in overcoming obstacles" (C 27:250). In a sense, the Stoics already saw beyond their eudaimonism, and glimpsed the supernatural ideal: "The sect of Zeno, like that of the Platonists, proceeded solely from the principle of morality . . . hence the maxim, to wrap ourselves in virtue, as if in a cloak. But to abstract from all enjoyment is contrary to nature, since there are so many needs that call for satisfaction" (V 27:484).

The Platonic ethical ideal takes this next step explicitly: it is "a mystical ideal, in which the highest good consists in the human being seeing himself in communion with the highest being" (C 27:250). What Kant refers to as the "Platonic" ideal seems more to be associated with the philosophy of the

later Academy or even with neo-Platonism. "Plato [he says] derived all sources of our understanding from God, and said that all our happiness would consist in raising ourselves above the sensible and uniting ourselves with our source of all in the highest being" (C 27:105).

The highest ideal of antiquity, however, Kant locates in the Christian ideal of holiness. The transition to this from the Platonic ideal consists in the thought that through communion with the highest being, we may be open to the influence of this being on our life and character. "Plato, in particular, apart from the principle of morality he derived from the power of reason in the human being, assumed also a mystical principle, which he located in the influence of a supreme being on the human mind" (V 27:484). The highest point in the ethics of antiquity, as Kant sees it, is the Christian ideal of holiness. The chief superiority, however, consists for him in the consequent Christian doctrine that the moral ideal is too lofty for human beings, and that they must therefore depend on divine assistance in order to reach the ideal.

It is this alone that releases the Christian ideal from what Kant thinks is the moral corruption attendant upon ideals in general. For ideals are drawn from human models, and are therefore always attended by deception and compromise, owing to our corrupt tendency to demand no more of ourselves than what we find in the imperfect human conduct around us. "The ideal of holiness, as philosophy understands it, is the most perfect ideal, for it is an ideal of the greatest purely moral perfection, but such a thing is unattainable by the human being, it is based upon a belief in divine assistance" (C 27:251f; V 27:485). All other ancient ideals, Kant argues, remained imperfect, because they "had no greater moral perfection than that which could come from human nature; but since this was very defective, their moral laws were also defective. So their ethical system was not pure; they accommodated virtue to human weakness, and hence it was incomplete" (C 27:251).

6 From ideals to principles

The superhuman and supernatural perfection of the ideal of holiness in the Gospel is for Kant the historical route that led to the modern conception of ethics as based on a *principle* rather than an *ideal*. For human ideals are always based on human models; a true model of perfection must therefore rise above what can be given to us as a model for imitation: "All ideals are fictions. We attempt *in concreto* to envisage a being that is congruent with the idea. In the ideal we turn ideas into a model, and may go astray in clinging to an ideal, since it can often be defective" (M II 29:605).

Every ideal, when its rational basis is understood, depends on something more fundamental, an idea of reason, which is generated through a search for what is unconditioned or perfect. The defect in every moral ideal is not only that it is represented as something empirically existing, which would always be defective, but even more fundamentally that it presupposes, and yet at the same time also conceals, the activity of reason which leads to the concept of perfection that the ideal is supposed to represent. To see to the bottom of the ethical ideal generally is to see that ethics cannot ultimately rest on ideals but on that which serves to criticize every ideal and is presupposed by it as a valid ideal. The moral *ideal*, therefore, must always rest on a moral *principle*. "The ideal is the prototypon of morality. A natural human being can never be the ideal, for he is still subject to weakness. The ancients would seem to have exhausted all possibilities here. But if we ask: "What is moral perfection, and on what principles is it to be judged? Then we can and must enter on new paths at this point" (M II 29:605). Modern ethics is therefore deeper than ancient ethics, both because it sees through the corruption of ideals and because it perceives their basis in rational principles.

Kant remarks in his logic lectures that in moral philosophy we have not come further than the ancients (L 9:32). If this is to be consistent with his portrayal of the history of ethics in the ethics lectures, he must mean by this only that the ancient ideals include everything belonging to the *content* of morality, not that there is no significant difference between ancient and modern moral philosophy, or that modern moral philosophy has taken no decisive step beyond antiquity. For it is clearly Kant's view that in *form*, modern moral philosophy certainly does represent a significant advance over ancient philosophy. It has, namely, transformed moral philosophy from a theory of ideals into a theory of principles.⁴

This, then, is the authentic Kantian reply to the recent criticisms of modern moral philosophy coming from those that call themselves representatives of "virtue ethics." The reply is historical, and its substance is that the style of moral thinking favored by "virtue ethics" was well suited to antiquity but can now be seen as naïve, shallow, and no longer appropriate

⁴ In context, what Kant is saying is that although it is fairly obvious that modern ethics has made no substantive advances beyond ancient philosophy, it might seem that modern metaphysics has advanced beyond that of the ancients. His point is that in matters of substance, however, this is not true: the apparent advances are all illusory. Modern metaphysics can be said to have advanced beyond the ancients only in form, by undertaking a critique of our capacities for metaphysical knowledge (L 9:33f). The remark, taken in context, is clearly compatible with the claim that in ethics too, there have been modern advances, but only in form and not in content.

to the modern world, in which we have come to recognize both the unreliability and imperfection of ideals and models, and the truth that every ethical ideal must be grounded on a principle which guides us in judging any model proposed for our imitation. Kant's own criticisms of virtue ethics are therefore continuous with those of some more recent Kantians, who claim that an ethics of virtues or ideals (or "ethical being") is incapable of dealing with such phenomena of modern society as cultural pluralism, social change, and the need to communicate rationally about what to do between people who have been formed according to very different cultural schemes of moral education.⁵

7 Kant's anti-Pelagianism

Kant credits the Christian ideal of holiness, because it represented the ideal as something superhuman and supernatural, attainable only through divine assistance, with being the historical means through which this crucial insight entered into the history of ethics. "In the Gospel everything is complete, and there we find the greatest purity and the greatest happiness. The principles of morality are presented in all their holiness, and now the command is: you are to be holy; but because man is imperfect, this ideal has an adjunct, namely divine assistance" (C 27:252).

This turn in Kant's thinking may come to us as a surprise or even a shock. Kantian ethics is supposed to be a theory of autonomy, a theory that encourages human beings to govern their own lives through reason and think for themselves. Yet here he seems to be opting for *Schwärmerei* over critique, theological morality over rational morality, the moral passivity he regularly condemns in Pietism over the moral autonomy on which his own opposed theory is based. Kant seems to be endorsing the Christian (and the specifically Pauline, Augustinian, and Lutheran) doctrine that the true morality is one that regards human agency as morally impotent unless assisted by divine grace. Our aim, on this view, should apparently not be human morality or endless progress, but superhuman holiness; moral ideals that depend on our natural powers are misguided and even corrupt; and we are in a state of total depravity unless we are given help from above. Kant seems to be siding with Augustine against the heresy of Pelagius, whose name is universally hated throughout all Christendom merely because he maintained the reasonable and even self-evidently correct position that we should be morally required to do only what lies within

⁵ See, for instance, Schneewind 1997b; Loudon 1997; Loudon 1992; and Wood 1999, 331–333, 416f.

our power and should be given moral credit or blame only for what we ourselves have done.⁶

There are indeed a few Christian scholars of Kantian ethics even today who think he should be read in something like orthodox Augustinian terms.⁷ But theirs is clearly a minority position in Kant scholarship, and seems motivated more by religious faith (with its characteristically corrupting influence on intellectual honesty) than on a sound reading of the texts. We might also think that the remarks just quoted could be best explained by the fact that they are found in Kant's lecture transcriptions, in which he is arguably catering to the religious beliefs of his audience, which were surely more in line than his own with Christian orthodoxy. My suggestion, however, will be different from both these. I propose to understand Kant's remarks as directed solely to the subject matter that they are explicitly supposed to be about, namely, the *history* of ethics.

There are three principal Kantian claims about ethics, whose progressive recognition in the course of the history of ethics Kant is attempting to present, and also explain:

- a. The ethical ideal is inadequately grasped when the highest good is conceived as happiness. Call this Kant's *antieudaimonism*.
- b. In formulating ethical ideals, human beings are in danger of confusing what ought to be with what they see around them, adapting the requirements of morality to human weaknesses. Call this Kant's thesis of the *imperfection of ideals*.
- c. The foundations of ethics are inadequately conceived when they are formulated as an *ideal* to be imitated; instead, these foundations must be conceived as a moral *principle*, which is required even for the proper formulation and criticism of any ethical ideal. Call this Kant's rejection of an ethics of ideals in favor of an *ethics of principles*.

⁶ "Where theism forms the principle of any popular religion, that tenet is so conformable to reason that philosophy is apt to incorporate itself with such a system of theology . . . [Yet] philosophy will soon find herself very unequally yoked with her new associate; and instead of regulating each new principle, as they advance together, she is at every turn perverted to serve the purposes of superstition . . . Amazement must of necessity be raised; Mystery affected: Darkness and obscurity sought after: And a foundation of merit afforded to the devout votaries, who desire an opportunity of subduing their rebellious reason, by the belief in the most unintelligible sophisms. Ecclesiastical history sufficiently confirms these reflections. When a controversy is started, some people always pretend with certainty to foretell the issue. Whichever opinion, say they, is most contrary to plain sense is sure to prevail; even where the general interest of the system requires not that decision. Though the reproach of heresy may, for some time be bandied about among the disputants, it always rests at last on the side of reason" (Hume 1956 [1757], 53f).

⁷ See, for example, Hare 1996; Palmquist 2000.

In the history of ethics, Kant thinks that the first form of antieudaimonism was the adoption of a supernatural ideal (whether Platonic or Christian) in place of a natural ideal (of Cynicism, Epicureanism, or Stoicism). The first recognition of the imperfection of ideals was the Christian ideal of holiness, and the attendant doctrine that human morality is possible only through divine assistance. The Christian ideal thereby also paved the way for the decisively modern turn in ethics, away from an ethics of ideals and toward an ethics of principles. In this way, Kant belongs to a central tradition in thinking about the history of Western culture, which in one way or another regards Christianity as having initiated a revolution in thinking that, when it eventually came to maturity, brought about a decisive separation of the worldview of classical antiquity from the worldview of modernity. These *historical* theses might be true even if there is no truth at all in the claims that the right ethical ideal is supernatural, or that the right ideal is that of holiness, or that human morality is possible only through God's supernatural help.

Kant appears to accept the Augustinian position to this extent, that our innate propensity to evil seems to stand in the way of the possibility that we might reform ourselves through our own effort: "But does not the thesis of the innate corruption of the human being with respect to all that is good stand in direct opposition to this restoration through one's own effort? Of course it does, so far as the comprehensibility of, i.e. our *insight* into, its possibility is concerned" (RGV 6:50). But it does not stand in the way of our assumption that reform is possible. "For if the moral law commands that we *ought* to be better human beings now, it inescapably follows that we must *be capable* of being better human beings" (RGV 6:50).

Kant does not dogmatically deny the doctrine that divine assistance is required for our moral conversion. "Everyone must do as much as it is in his powers to do; and only then, if a human being has not buried his innate talent (*Luke* 19:12–16), if he has made use of the original predisposition to the good in order to become a better human being, can he hope that what does not lie in his power may be made good by co-operation from above" (RGV 6:52). Kant emphasizes that nothing but good conduct can be regarded as a pre-condition for receiving this aid – to think of some irrational belief state, for instance, or ceremonial expressions of penitence or groveling before the divine being as such conditions would be "religious delusion," "fetishism" and "superstitious counterfeit service" of the Deity (RGV 6:190–200). "It is not essential, and hence not necessary, that every human being know what God does, or has done, for his salvation; but it is essential to know *what a human being has to do himself* in order to become worthy of this assistance" (RGV 6:52, cf. 6:171f).

Kant's actual position on the Augustinian-Pelagian controversy, then, is a rather subtle one, but also a deeply troubled one, in several respects. This position is not obviously consistent with his own doctrines, and it is clearly a position that neither orthodox Christians nor many present day Kantians (at least those, like myself, who are more naturalistic, and not orthodox Christians) could comfortably accept.

8 Conclusion

Kant was, then, after all, a historian of philosophy, even a historian of ethics. More importantly, he saw his own views in ethics in a determinate historical perspective. In treating the history of ethics, moreover, he used inventively many of the same devices employed by other more famous philosophical historians of philosophy, such as Aristotle and Hegel. He saw past views in light of the range of options open to a philosopher, understood contrasting views as one-sided or incomplete, and presented his own view as mediating them or transcending their common defects; he portrayed the history of ethics as a narrative embodying a progression. Whether or not we find Kant's historical narrative persuasive, either as historiography or as philosophy, by appreciating it we gain better grasp of Kant's own historical self-understanding as a moral philosopher.

*Moral obligation and free will**Oliver Sensen***Introduction**

What is the defining feature of morality? When does something fall in the moral domain in contrast to not being morally relevant? For instance, there might be rules of etiquette or legal norms which as such might not state a moral demand. What is the distinguishing feature that makes a rule a *moral* norm? It is unlikely that the answer will be satisfactory if one points to a particular content that all moral norms are supposed to share. If, for instance, one were to say that moral norms concern the rights of others, this might be too wide and too narrow at the same time. Other subjects might be concerned with the rights of others, e.g., legal history, there might be moral rules which do not concern rights, e.g., rules of holding open doors for people coming behind, and there might be rules that do not concern others, e.g., there might be duties toward self.¹

Kant's answer is a formal one.² The central feature of morality, according to Kant, is an unconditional ought, that one should do something independently of what one desires. It has been argued that such a view of morality is deeply engrained in human beings, so that already children of a young age regard moral rules to be valid universally and independently of what one wants (cf. Nichols 2004, 5–7). And, paradoxically, even a sentimentalist who argues that morality is based on desires, would also argue that – when it comes to acting morally or not selfishly – one should not have one's desires in mind (cf. Butler 1729, Sermon XI; Feinberg 2008, 522f). In Kant's works the central place of the unconditional ought, or obligation, can easily be overlooked if one just takes his published works into account. But this central role becomes clear in Kant's *Lectures on Ethics*. As I shall argue, this is not just because Kant used Baumgarten's *Initia*, in

¹ On Kant's account of such duties, see Lara Denis's contribution to this volume.

² For the difference, see Wolf 2013, 42.

which obligation is the key concept,³ as the textbook for the first part of the *Lectures*.

In this chapter I shall first demonstrate the importance Kant accords to obligation (Section 1), I shall then analyze the concept as Kant uses it (Section 2), and his arguments for why obligation is a central concept (Section 3), before discussing whether there really is such a thing as unconditional obligation (Section 4).

1 The importance of obligation in Kant's works

The term “obligation” [*Verbindlichkeit*] is not one of the first terms one would think of in summarizing Kant's moral philosophy. Conceptions such as the good will, the categorical imperative, respect for humanity, autonomy or the highest good are much more likely candidates that come to mind. Yet obligation is at the heart of Kant's moral philosophy, and the link behind many of the other concepts.

For instance, obligation is one of the “prior concepts”⁴ to the *Metaphysics of Morals* (cf. MS 6:222). In explaining the foundation of his moral philosophy Kant refers to the concepts of freedom, law and obligation (cf. MS 6:221–226). Similarly, although the term “obligation” hardly appears in the *Groundwork* at all (cf. G 4:389, 439), it does play a central role there too in its close relationship to duty. While “obligation” refers to the bindingness of morality as such, “duty” expresses the obligation to a particular action: “The objective necessity of an action from obligation is called *duty*” (G 4:439, cf. 400).⁵ Duty is of prime importance in the *Groundwork* for two reasons. First, it adds the specific quality that distinguishes mere law-full behavior from morally worthy actions: “For, in the case of what is to be morally good it is not enough that it *conform* with the moral law but it must also be done *for the sake of the law*” (G 4:390). This cast of mind that is required in addition is acting out of a sense of duty or respect for the law: “*duty is the necessity of an action from respect for law*” (G 4:400). Second, Kant uses this requirement of duty to derive or formulate the content of what the moral law says (cf. G 4:402, 421). The key feature of the central concept of duty is obligation.

³ Cf. Schwaiger 2009, 68; for Kant's relation to Baumgarten, see Bacin's contribution to this volume.

⁴ This seems a better rendering of “*Vorbegriffe*” than Gregor's “preliminary concepts” (MS 6:221).

⁵ Kant's usage of “duty” is different in the *Metaphysics of Morals*, where it just refers to the action that is commanded: “*Duty* is that action to which someone is bound” (MS 6:222).

The importance of obligation in Kant's moral philosophy can also be demonstrated historically. In his earliest writing on the subject, Kant already laments, "how little even the fundamental concept of *obligation* is yet known" (*Inquiry concerning the distinctness of the principles of natural theology and morality* AA 2:298). From then on he repeatedly argues in his lectures that alternative theories to his own cannot yield obligation. The criticism changes little over time (cf. H 27:9; P 27:106–110, 118–120; C 27:252–255, 275–278; M II 29:621–629; V 27:497–499). But it is not before Kant wrote the *Groundwork* that he has fully developed his own solution to the problem of obligation. In his answer Kant proposes a special source of the moral imperative. The point of the imperative is to express an unconditional ought: "The categorical imperative, which as such only affirms what obligation is, is: act upon a maxim that can also hold as a universal law" (MS 6:225). Unconditional obligation becomes possible if the will gives itself its own law, or has autonomy: "*Autonomy* of the will is the sole principle of all moral laws and of duties in keeping with them; *heteronomy* of choice, on the other hand, not only does not ground any obligation at all but is instead opposed to the principle of obligation" (KpV 5:33).

In sum: although the term "obligation" is not the first concept that comes to mind if one thinks about Kant's moral writings, it is at the heart of his moral inquiry. The question of how unconditional bindingness is possible has occupied him from the beginning of his writings on moral philosophy, and his answer to the problem of obligation is his conception of autonomy (see later in this chapter). The categorical imperative, autonomy, and duty all describe aspects of obligation. Kant is much more explicit about obligation in the *Lectures*, and they will serve to clarify how Kant conceives of obligation.

2 Kant's conception of obligation

What, then, is obligation, according to Kant? "To have an obligation anywhere, or be bound to something" (V 27:493) expresses "an *ought* or a *necessitation*" (V 27:488). To stand under an obligation is therefore to be necessitated in some way. Necessitation [*Nötigung*] is a "making necessary" (M II 29:611; C 27:255f): "Necessitation is an *actio* whereby a thing is made necessary, that was not so before" (M II 29:605). There are different ways in which something can be made necessary. For instance, natural laws determine what "comes about in a necessary manner" (V 27:488). In the realm of action a behavior might be made necessary by a stimulus, that is, a "cause that determines man . . . according to the laws of nature" (V 27:493). This

might be the case for most of the animals: “animals have no free choice, their actions being necessarily determined by their sensory impulses” (M II 29:611). Kant calls this kind of making necessary a “pathological necessitation.” If humans are not to be wholly determined this way, “the necessitation must be practical, or *per motiva*.” Motiva are not sensible stimuli, but are “representations of the understanding, and of reason, that determine the will” (M II 29:611). There are two kinds of *motiva*, depending on whether the representations of reason relate to our well-being or the moral worth of our person. Kant calls the first kind “pragmatic,” and the second kind “moral motives” (M II 29:612). Pragmatic motives necessitate only conditionally, while moral motives do so unconditionally. Kant reserves “obligation” for moral necessitation: “*Necessitatio moralis* is *obligatio*” (M II 29:612).

The kind of necessitation Kant is interested in when he talks about obligation is therefore different from a predetermination of actions, as would be the case if our actions were wholly caused by stimuli: “Necessity and necessitation are different” (M II 29:611). Necessitation expresses that something ought to be the case, not that it will happen unavoidably: “An action is necessary if one cannot resist it” (C 27:267). Obligation is not necessary in that sense. Conversely, this means that a being who could not act other than in a morally good way, would know no obligation: “With God, the objective practical law is also, at the same time, a motive” (M II 29:605). It is therefore “impossible, rather, that He should will anything else” (M II 29:607). God would necessarily act in the right way, and for the right motive, since he does not have any inclinations: “Hence practical necessity, in the eye of the supreme being, is no obligation; He necessarily acts morally, but has no obligation” (C 27:256).⁶

Obligation therefore involves a certain form of compulsion [*Zwang*]: “All obligation is a kind of compulsion” (C 27:269). Only a being who might be tempted to transgress morality can be under an obligation: “A compulsion always presupposes a hindrance in the will” (M II 29:616). In the case of human beings these hindrances are sensible impulses and desires: “Self-seeking always subordinates him to the worth of his total condition, but reason thinks otherwise” (M II 29:612). Obligation therefore comes about in human beings because there is a conflict. One’s desires and the moral law do not always pull in the same direction. The compulsion is felt if one acts morally, albeit reluctantly. Compulsion “consists in the necessitation to an

⁶ Robert Stern emphasizes this side of obligation, cf. his 2012, 41–99.

action that he undertakes with reluctance . . . A thing is done reluctantly by a free being, insofar as (1) there is present in him an inclination to the opposite of what he *wills* to do and (2) he nevertheless does what he *wills* as a free being" (V 27:519). The moral law therefore appears to human beings as an imperative (cf. G 4:413, KpV 5:20).

So what, more specifically, is obligation? For instance, is it a feeling of being coerced [*genötigt*], is it the fact that an action is declared to be necessary, a mix of the two, or something else altogether? One way of relating Kant's thought to our ordinary experience might be to say that it relates to what we ordinarily might call a bad conscience, a feeling of being constrained to an action. What speaks in favor of identifying obligation with a feeling is that Kant himself says that a moral feeling would be among "*subjective* conditions of receptiveness of the concept of duty" and that it is "by virtue of them that he can be put under obligation" (MS 6:399). Moral feeling, more specifically, is "the susceptibility to feel pleasure or displeasure merely from being aware that our actions are consistent with or contrary to the law of duty" (MS 6:399). Without such a feeling we would be "morally dead" (MS 6:400), since "any consciousness of obligation depends upon moral feeling to make us aware of the constraint present in the thought of duty" (MS 6:399). So, does this mean that Kant talks about a feeling when he refers to obligation?

This interpretation seems to be contradicted by other statements Kant makes. For instance, he says that if feeling were the ground of morality, then "not everyone would have to be obligated in the same degree, for not everyone has the same feeling" (M II 29:625). How could such a feeling ground duty, where "duty is by its nature absolute, unconditioned and necessary" (V 27:506)? Accordingly, Kant states that "this constraint (*necessitatio*) is the determination of the human will . . . through moral rules" (V 27:485). The moral rule is "the sufficient reason for determining" the will (V 27:487), and "the action . . . is made necessary by the moral law" (V 27:489). This suggests that obligation is the fact that the moral law declares something to be necessary. One knows the binding character of the law even *before* the moral feeling: "the agent must have knowledge of the law and its binding character, before he can be filled with pleasure or pain on obeying or violating it. . . . The *moral feeling*, . . . is . . . not the ground of duty." (V 27:498, cf. 506).

How do these two different statements – about the importance of feelings and the moral law – go together? Kant says that the role of moral feelings is as one of the "*subjective* conditions" of the consciousness of obligation, "not as objective conditions of morality" (6:399). This suggests that obligation is

a complex phenomenon that has more than one facet. Being obligated is the fact that the moral law declares an action as necessary. This declaration then has an *effect* on feeling (cf. KpV 5:74f). It is “the authority of the laws, which enjoins absolute obedience, and awakens resistance and struggle” (V 27:490). This is a feeling component that we recognize in everyday experience, and if we never had this experience, we would be morally dead. But the central feature of obligation as such is not the feeling. There could, as we also know from common experience, be such a thing as an erroneous conscience, or feelings that present themselves as absolute (cf. Blackburn 1993, 177), which in fact do not make up a strict obligation. The hybrid interpretation of Kant is confirmed by the following statement.

By categorical imperatives certain actions are *permitted* or *forbidden*, . . . while some of them or their opposites are morally necessary, that is, obligatory. For those actions, then, there arises the concept of a duty, observance or transgression of which is indeed connected with a pleasure or displeasure of a distinctive kind (moral *feeling*), although in practical laws of reason we take no account of these feelings (since they have nothing to do with the *basis* of practical laws but only with the subjective *effect* in the mind when our choice is determined by them, which can differ from one subject to another [without objectively, i.e., in the judgment of reason, at all adding to or detracting from the validity or influence of these laws]). (MS 6:221)

Obligation is foremost the fact that the law declares something as necessary, but in human beings this declaration has an effect on feelings.⁷ The central role the moral law plays in obligation is confirmed by Kant’s discussion of traditional distinctions within obligation. For instance, there seems to be a difference between being obligated (passive obligation), and obligating someone (active obligation): “All obligation is distinguished as active or passive” (M II 29:613; trans. Timmermann). However, Kant deems this distinction to be “unimportant” (C 27:260). The reason is that even if I oblige someone, he is only bound to the action because of the moral law which is in him. In this sense even duties toward others depend on duties toward self (cf. also MS 6:417f). All obligation follows from the moral law, and therefore there is only passive obligation: “In actual fact there is only passive obligation” (M II 29:613; trans. Timmermann). If another coerces me, he needs to refer to the moral law: “I can be coerced, for example, by others into payment of debt, albeit only through the idea of binding law” (V 27:523). But this law is accessible to me by my own reason:

⁷ On the role of feelings in Kant’s moral philosophy, cf. Sensen 2012b.

the other, having a right to do so, confronts the subject with his duty, i.e., the moral law by which he ought to act. If this confrontation makes an impression on the agent, he determines his will by an Idea of reason, creates through his reason that conception of his duty which already lay previously within him, and is only quickened by the other, and determines himself according to the moral law. (V 27:521)

In claiming a right (active obligation), I therefore remind the agent of his own duty to obey the law of his own reason (passive obligation, cf. MS 6:239). With this understanding of the source of obligation, Kant sees himself as doing something new, in direct contrast to his predecessors. Prior to Kant, philosophers put the source of obligation in the will of another, for example, a ruler or God (cf. V 27:510; C 27:262; Schneewind 1998, 518–522). But, according to Kant, this would turn all obligation into a form of positive law, where the laws “have their ground in the choice of another” (V 27:510), in contrast to being based on natural laws. (I shall address Kant’s argument for this claim in the [next section](#) of this chapter.) Kant himself sees moral obligation as based on a natural law, where “obligation can be derived and known solely from the nature of the action by laws of freedom” (V 27:510). In a figurative way, Kant can uphold the traditional view that obligation involves two, the obligator and the one who is obliged. But this other person would be one’s own reason as the source of the moral laws: “if the obligator is personified as an ideal being or moral person, it can be none other than the legislation of reason” (V 27:510). Obligation arises between two aspects of the same human being: “this, then is man considered solely as an intelligible being, who here obligates man as a sensory being, and we thus have a relationship of man *qua* phenomenon towards himself *qua* noumenon” (V 27:510). However, this is not meant as a claim about two metaphysically distinct entities, a noumenal and a phenomenal being. Rather, the noumenal being is just the idea a person has of himself as being morally good in accordance with the moral law:

It seems as though, in duty, the will of a legislator underlies, not anything we do by our own will, but what we do by the will of another. Yet this other will is not that of another being; it is only our own will, insofar as we make it general, and regard it as a universal rule. Such a will operates as a universal, not as a private will. (M II 29:627, similarly MS 6:419)

This basic source of obligation, the moral law of one’s own reason, is not different in the case of legal obligation: “The difference between law and ethics does not consist in the kind of obligation, but in the grounds that motivate us to fulfil obligations” (C 27:271). Ethics demands not only that

one do the right thing, but also that one do it simply because it is right (cf. again G 4:390). In contrast, law merely demands the outward performance of an action, for example, to stop at a red traffic light, independently of what one's motives are. But the obligation to perform or refrain from a course of action rests in both cases on the same law:

Within this universal moral law are comprehended both legal and ethical laws, though with this difference, that in the former actions are considered merely in regard to their form, whereas in the other they are considered with respect to their end, as their object. (V 27:526)

In sum: the central feature that makes moral obligation possible, according to Kant, is the moral law of one's own reason. Obligation is therefore necessitation by the moral law: "Necessitation by the moral law, to act in accordance with it, is *obligation*" (V 27:481). Obligation is thereby "the dependence of choice on a necessitating law as a proximate cause" (V 27:492). But why does Kant think that obligation must rest on one's own reason? Why is this a condition for the possibility of obligation?

3 Autonomy as condition of moral obligation

Kant argues that only autonomy makes obligation possible: "*heteronomy* of choice, on the other hand, not only does not ground any obligation at all but is instead opposed to the principle of obligation" (KpV 5:33). Autonomy is the "law-giving of human reason" (V 27:499). As such, autonomy refers to the source of the moral principle. The moral principle is not imposed from outside, but originates in one's own reason. By this Kant cannot mean the deliberate adoption of a principle by an agent, such as a New Year's resolution. For in that case "the one imposing obligation . . . could always release the one put under obligation . . . from the obligation . . . , so that (if both are one and the same subject) he would not be bound at all to a duty he lays upon himself" (MS 6:417). Instead, the law-giving Kant talks about is the one of pure reason, that is, it is prior to one's awareness and conscious choosing. One could compare it to the principle of noncontradiction in theoretical reasoning. If one reasons, one is guided by this principle, although one can still act against it. Similarly, the moral law is a constitutive principle of pure reason that necessarily guides moral reasoning.⁸

Kant's point – to give a preview of his argument – is that if the source of the moral law is anywhere else than in the autonomy of one's own reason,

⁸ For a fuller defense of this view, cf. Sensen 2013; for his notion of autonomy, see Sensen 2012a.

then one would need an incentive to comply with that law. First, this would be heteronomy because one would be guided by the natural laws of the incentives, and not by one's own law. Second, it would undermine obligation because we hold moral obligation to be universal and necessary, but incentives are relative and contingent, and could not ground something necessary and universal:

Thus if, for example, the principle of universal happiness were to be the basis for determination of the moral laws, it would be a question how far our needs were satisfied in their entire totality by following these laws; but here laws of nature are involved . . . so that reason would have to obey the laws of nature and sensibility . . . But this would obviously put an end to the autonomy of reason, and thus be heteronomy. (V 27:499)

If one acts to satisfy a need or desire, one is governed by the laws of the desires. The problem with the dependence upon these is that desires are relative and contingent, and cannot ground necessary and universal moral laws: "But that the determining grounds of laws of duty cannot be built upon natural laws is already evidenced by the quality of these laws of duty, namely that they must be necessary and universally valid" (V 27:499).⁹ The laws to procure happiness are not necessary and universal since there is no universal and definite concept of happiness (cf. G 4:418), and "the means to happiness are too diverse to be determinable in advance" (V 27:500).

Kant supports his claim that the categorical imperative is the right expression of morality by arguing that all alternative theories to his own would yield heteronomy. If this is so, then none of these theories could yield moral obligation, and only his account would be left. His exact argument why all other theories yield obligation is in need of some reconstruction. For instance, Kant says that all alternative theories yield heteronomy, but then criticizes the command "be perfect" as being by itself empty (cf. G 4:443). Why would acting according to that command be heteronomy then? The *Lectures* are helpful in reconstructing Kant's arguments. They confirm the picture of the published writings, but are often more elaborate and explicit. In addition, the *Lectures* would allow reconstructing the development of Kant's criticisms, as it is only since the time of the *Groundwork* that Kant connects his objections with the concept of heteronomy. The *Lectures* therefore help to bring out what is essential to Kant in his objections toward alternative theories. In the following, I shall cite the lectures where possible.

⁹ That morality must be necessary and universal is, according to Kant, "the generally received concept of morality" (G 4:445, cf. 389). But Kant does not rest his case on a common notion, he will later argue that there really is such a morality (cf. [Section 4](#)).

The Mrongovius lecture notes in particular give a very rich account of Kant's objections.¹⁰ As in his published writings (cf. G 4:440; KpV 5:40, cf. Schneewind 2009), Kant divides alternative grounds for morality into empirical or rational bases (cf. M II 29: 621).¹¹ Within these he distinguishes again between internal and external grounds. Internal empirical grounds of morality are physical feeling (I.1.1) and moral sense (I.1.2), external empirical ones would be custom (I.2.1) and education (I.2.2). The internal rational ground Kant repeatedly discusses is perfectionism (2.1), the external one the will of an arbitrary God (2.2). Kant does not only regard these as "all previous," but also as "all possible" (KpV 5:39; G 4:432, 441) alternative theories in that he regards the distinction to be exhaustive. (I shall discuss whether there are in fact more alternatives below.)

Kant subsumes all these other theories under heteronomy: "The principles aforementioned are principles of heteronomy" (M II 29:629). But why? For some of them it will be easily granted that they depend on foreign laws (of society or an arbitrary God), but why could they then not yield obligation? For others, one's own feelings and the rational idea of perfection, one might wonder whether they depend on *foreign* laws at all. Kant's answer to these worries is that all the possibilities need an inclination in order for one to adopt their principles. Regardless of the source of these prescriptions, they all would need a desire as their determining ground, i.e., their "motivating cause" (V 27:493). If that is so, then all other theories than his own are dependent on the natural laws that govern one's desires, and therefore yield heteronomy. But why does he think that all of them are dependent upon desires?

The least controversial case regarding the dependence on desires will be I.1.1, moral theories built on physical feeling. Such theories are by design based on desires.¹² For Kant, the principle that is based on physical feeling "is that of self-love, and rests upon the comfort and safety of our condition" (M II 29:621), or, in short, one's happiness: "the state of happiness consists in consciousness of enjoying and possessing the means to procure for ourselves all ends that are even possible, and thereby to satisfy all wishes" (V 27:497). But Kant also subsumes theories built on a moral sense (I.1.2) under happiness or the physical feeling: "I count the principle of moral feeling under that of happiness because every empirical interest promises to

¹⁰ On these notes, see Jens Timmermann's contribution to this volume.

¹¹ Kant sometimes labels the same distinction "subjective" and "objective" (cf. KpV 5:40).

¹² For our present purposes we can treat feelings and desires interchangeably. On Kant's definitions, see Sensen 2012b, 46–48.

contribute to our well-being by the agreeableness that something affords, whether this happens immediately and without a view to advantage or with regard for it" (G 4:442n.). A feeling of assent or sympathy would make a positive contribution to one's own well-being. Kant argues that all feelings must be of one kind so that one could outweigh the other:

At bottom we have only one feeling, namely pleasure and pain, and this is the judgement upon our overall well-being. There are various kinds of sense, but only one feeling of pleasure. If there were several feelings, or a power of distinguishing by satisfaction, we could not distinguish feeling by degree. We compare with one another the pleasure at a fine speech and the pleasure to eat a dish. (M II 29:625; cf. KpV 5:23)

On Kant's account, all internal empirical grounds therefore rest on desires. But Kant argues that internal *rational* grounds (2.1) are also based on desires. He mostly emphasizes that a command "perfect thyself" would be "an empty husk" (M II 29:626). It would be a tautology because one still would need to know in what the perfection lies: "For if we want to know what the ground of this perfection is, and receive as an answer: Seek perfection, then it all comes to the same" (M II 29:626). However, this by itself does not explain why the command would yield heteronomy. Why does the command depend on a foreign law? Kant's answer is that to perfect oneself could only mean to develop one's talents, and one would ultimately develop them for one's happiness, and therefore perfection would be based on the desire for happiness as well: "This perfection, as a *characteristic* of the human being and so as internal is nothing other than *talent*," yet "talents and their development . . . can become motives of the will only by means of the happiness we expect from them" (KpV 5:41). So, in order to adopt perfection, it would be one's desire for happiness that would be the determining ground. Since desires depend on natural laws, perfection as the moral principle would be heteronomy too.

For the two remaining cases, morality based on the will of an arbitrary God (2.2) or customs of society (1.2), it seems clear why they would ground morality on laws external to the agent. But this is not the law Kant has in mind. Rather he also reduces those options to subjective desires and their natural laws. If voluntarism were right (2.2), and God does not command moral laws because they are right, but if moral laws would be right because God commanded them at will (and could as easily command that murder is right, for instance), then this would also yield heteronomy. The reason is that one would need a desire to please God or not be punished by him in order to adopt his commands:

Supposing the *arbitrium* of God to be known to me, where is the necessity that I should do it, if I have not already derived the obligation from the nature of the case? God will it – why should I? He will punish me; in that case it is injurious, but not in itself wicked; that is how we obey a despot; in that case the act is no sin, in the strict sense, but politically imprudent. (H 27:9)

Arbitrary external commands would need desires as their determining ground: “Actions, in that case, will be based upon fear and hope” (M II 29:627f).

The same problem arises for external empirical grounds of morality, that is, if morality were based solely on custom (I.2.1) or education (I.2.2). If society prescribes laws, and adds “authority, which coerces us thereto with punishments” (M II 29:621), one would obey these commands out of fear or hope. Again, desires would be the determining ground, and one’s ground would be heteronomy, since it depends on the natural laws of one’s desires. (This does not mean that there cannot be any political obligation: cases where a state introduces an obligation to drive on the right side of the street, for instance, “but then a natural law would still have to precede it, which would establish the authority of the lawgiver” [MS 6:224].)

So Kant rejects all the alternatives he considers because they would yield heteronomy. The problem with heteronomy is that it could not yield universal and unconditioned obligation. If morality rested on feeling, then there would not be a necessary and universal moral law. The law would not be universal because feelings differ from person to person and even in one and the same subject: “From the feeling of sensation that may be different in every creature, no generally valid law can be derived” (M II 29:625). More importantly in our context, a morality that needs feelings as the determining ground could not yield a necessary obligation:

Feeling, in man, is diversified, and that would also have to be so here. If morality rested on feeling, then many a one who is simply without tender feeling might attend to it less, and thereupon practice vice. If this were the principle of morality, then not everyone would have to be obligated in the same degree, for not everyone has the same feeling, and in degree the latter, in fact, is very varied. (M II 29:625)

Feelings differ, and if they were the reason why one acts, not everyone would be obligated in the same way. There would not be a necessary obligation. If Kant is right that all alternative theories depend upon desires for their determining ground, then nothing is left but Kant’s view that autonomy is the ground of a necessary obligation: “the question is: Insofar as it lies in reason, what is the action’s determining ground, from which

there arises a moral necessitation as the ground of obligation?” (V 27:495). The determining ground of heteronomy has been desires. In excluding these as unable to provide obligation, the only thing that remains is the form of the law (cf. G 4:402): “since the categorical imperative carries with it an unconditioned moral necessitation, which is founded not at all on the end or purpose of the action; so all that is left is the *form of lawfulness*” (V 27:495). Part of morally worthy actions is not merely a prohibition to adopt certain maxims, but also a motivational requirement, to do the right action simply because it is right (cf. once more G 4:390): “In other words, act so that you may present yourself, through the maxim of your action, as universally legislative . . . In this form of moral action lies, then, the determining ground of obligation, whereby it acquires the force of law” (V 27:496).

If so far Kant is right about his analysis, there still remains the question whether there are other possibilities he does not consider. Despite Kant’s claim that he has exhausted all possible alternatives, these are not the options we are working with today, and one can ask whether Kant’s objections would also affect our contemporary alternatives. Which are these? Shafer-Landau has proposed two sets of three-fold distinctions to cover moral theories (cf. his 2003, 13–8, 39–43). The first set of distinctions divides moral theories into (1) Nihilism, (2) Constructivism, and (3) Realism. Within Constructivism he then divides between (2.a) Subjectivism, (2.b) Relativism, and (2.c) Kantian ideal contract theories. Nihilism would be the view that there is no morality and no moral obligation. Realism would hold that there is morality, and that it exists independently of any human stance, whereas constructivism sees morality as constructed by human beings. This construction could depend on the stance of one person (2.a subjectivism), the actual agreement of many people (2.b relativism), or the construction that would happen by many under ideal circumstances (2.c Kantian theories).

How do these distinctions match with Kant’s? It seems that Kant would label subjectivism (2.a) as internal empirical (1.1.1), and relativism (2.b) as external empirical (1.2). If we assume that Kantian theories (2.c) cover his own,¹³ then this leaves nihilism (1), and realism (3) which Kant does not discuss at this point. By contrast, perfectionism in Kant’s account (2.1) does not appear in contemporary debates, and the voluntarism (2.2) he considers resembles more a relativism (2.b) than a contemporary realism (3). Would Kant have anything to say about contemporary forms of nihilism (1) and

¹³ For one reservation, see Sensen 2013, 63–66.

realism (3)? Nihilism by its own admission cannot account for unconditional obligation because it denies its existence, and at this point we are just concerned with the question of what could underpin unconditional obligation. This leaves just moral realism to consider. Kant does not address realism directly in the discussion of all “spurious” principles of morality (G 4:441). Realism was not among the previous moral theories Kant encountered.¹⁴ However, the arguments he gives in the paradox of method equally apply against a form of realism, according to which morality exists independently of any stance of reason. There Kant argues: “If the concept of the good is not to be derived from an antecedent practical law but, instead, is to serve as its basis, it can be only the concept of something whose existence promises pleasure and thus determines the causality of the subject, that is, the faculty of desire, to produce it” (KpV 5:58). This is the same thought as earlier.¹⁵ Even if the source of morality would be a reality independent of reason, one would also need a desire to act in accordance with that reality. This, too, would be heteronomy, and Kant subsumes the discussion of an independent moral reality under “all the errors of philosophers . . . their principle was in every case heteronomy and they had to come unavoidably upon empirical conditions for a moral law” (KpV 5:64).

Even if one follows Kant up to this point, so far he has only established a conditional: *If* there is unconditional obligation, it consists in following the categorical imperative for its own sake. But why should one think that there is unconditional obligation?

4 Freedom as the condition of moral obligation

At this stage of Kant’s argument, it is still possible that there really is no unconditional obligation, but that it is a mere “chimerical idea without any truth” (G 4:445). If there is unconditional obligation, it needs to have this particular form, Kant argues, but there might not be unconditional obligation. Could it be that Kant took an outdated religious model of obligation and transferred it unto morality (cf. Stratton-Lake 2012)? As in his published writings, Kant does not rest with the description of what obligation would be like, but he also asks: “But now how is a categorical necessitation to duty possible, and how can it be demonstrated?” (V 27:500). Here, too, Kant’s answer is that there is unconditional morality *if* human beings are free:

¹⁴ On the question whether Kant himself might be a moral realist, see Sensen 2013.

¹⁵ For a fuller treatment of this question, see Sensen 2011, 14–27.

“every being that is conscious of his freedom must also think that he necessitates himself to duty through the autonomy of his own reason” (V 27:501). In line with his previous work, Kant does not claim that he can spell out the exact mechanism of freedom, or prove theoretically that we are free, but he argues that we are justified in assuming our freedom.

But what exactly does Kant mean by freedom? He does not define it as the capacity to do otherwise. Nor does he contrast it with determinism: “every act of man must be determined, only with this difference, that if this ensues according to the law of free choice [the moral law], it must come about quite independently of all preceding circumstances” (V 27:502f). If an action would not be determined, it would not have a cause, and would be an accident, and could not be imputed to the agent. The contrast is therefore not whether an action is determined or has no cause, but whether it is determined in a previous time: “this must be called predeterminism, and not determinism” (V 27:503). Furthermore, in talking about freedom, Kant does not have an empirical reasoning in mind that weighs between two courses of action: “Man is not set free from the mechanism of nature by the fact that in his action he employs an *actus* of reason” (V 27:503). One’s reasoning as one becomes aware of it in introspection is still in time and might be predetermined by past influences or actions (cf. F 27:1322).

Instead, the freedom Kant has in mind is that someone is the first cause of an action without this cause being itself part of sensibility and predetermined in time. Imagine, for instance, that your desires incline you to give false testimony. If you refrain from doing so because the moral law prescribes that this would be morally wrong, you are the cause of action (to refrain from lying), without your action itself being determined by nature. Nature, your desires, inclined you to tell a lie, whereas your cause, the moral principle and its specific laws, lie a priori in reason: “The laws produce the causality in actions, i.e., the property whereby the agent becomes the cause of the action” (V 27:481). This conception of freedom as a first cause is – in the first instance – not the question of whether someone could have acted otherwise, but is more comparable to an unmoved mover conception. The unmoved mover is the first cause of a change, without himself being moved. This does not mean that the first mover could have acted otherwise than he did. In Kant’s case the mover moves itself by its own a priori law.¹⁶ The law of reason does not change, and it is in this sense that Kant says that freedom is outside time (cf. KrV A 539/B 568). The moral law commands to refrain

¹⁶ Huaping Lu-Adler suggested the term “self-moved mover” to me.

from giving false testimony, and if the agent is determined by it, he is the first cause of the action.

The ordinary sense of freedom as “could have acted otherwise” is secondary to this. If one is tempted by inclinations, the moral law makes it possible that one could act otherwise. But the paradigm of freedom for Kant is if one acts according to the moral law: “freedom, which *negatively* consists in the independence of choice from all determination *per stimulus*; . . . *positively*, however, it consists in spontaneity, or the ability to determine oneself by reason, without the need for triggers from nature” (V 27:494). If one acts morally, the action is not out of one’s control in the sense of being a product of natural laws, but can be imputed to oneself: “For this to happen, however, it is requisite that somebody can be regarded as the originator (*auctor*) of the action, i.e., as its complete first cause” (V 27:503). Freedom is being a (self-caused) first cause.

But how can there be such a freedom if nature is causally predetermined? In the *Lectures* too Kant does not claim to be able to explain how this is possible. As in his published writings, Kant merely tries to show that these two standpoints do not conflict, and that we are justified in regarding ourselves as free. The standpoint of sensibility is the standpoint of the observer, and the standpoint of freedom is the one of the moral agent (cf. V 27:504f). Therefore we do not declare ourselves to be predetermined and free in the same regard, we “view man from two sides, namely as phenomenon, i.e., as an appearance through his inner sense, and as noumenon, i.e., as he knows himself, in himself, through the moral laws” (V 27:505f). In addition to our natural being, we are also aware of the moral law that commands to act independently of nature.

How freedom can intervene in nature, we do not know: “Freedom cannot, therefore, be made comprehensible . . . ; only the belief that we are free is capable of explanation” (V 27:505). But why are we justified in regarding ourselves as free? We do not know that we are free, that is, capable of determining ourselves independently of nature, from experience, which always just shows us nature. Instead, we know freedom through “the categorical law of duty within us” (V 27:506). This is parallel to Kant’s solution in the second *Critique* (cf. KpV 5:29f): “Morality, therefore, is the sole means of obtaining consciousness of our freedom.” (V 27:506). How does this come about?

Kant gives an example of someone who is tempted to tell a lie. It involves “injury to my friend, physical pain, advantage that I may gain” (V 27:506), but at the same time one is aware by the moral law that one should not lie: “regardless of all evil, all physical force, there is a necessitation here to a

truthful testimony" (V 27:506). If one does refuse to give false testimony, the action is not caused by one's desires or sensible nature, but it is caused by the moral law which in itself is not caused by nature. This means that in being moral one is the first cause of the action, and thereby free:

I now determine myself through my reason; this is freedom, but this reason of mine is determined by the moral law, the very law that necessitates me to overcome the motives of nature. . . . There is thus within me a power to resist all sensory incentives, as soon as a categorical imperative speaks. (V 27:506f)

If one judges that it is wrong to give false testimony, one can gather from there that one can determine oneself freely: "The position, then, is that freedom is known by an inference (namely from the moral law) and not immediately felt" (V 27:507). But this does not mean that one can prove the existence of freedom, or that one knows it by experience: "For no man is in a position to determine in advance whether, *in casu dato*, he will . . . absolutely speak the truth" (V 27:507), nor – one can add – do we know after the fact from which motive we in fact acted (cf. G 4:406–408). But in judging that one ought to speak the truth, one is justified in assuming that freedom is possible. Ought implies can: "he knows only that he ought to obey the categorical imperative; he must therefore also be able to" (V 27:507).

Kant therefore goes further than merely stating the conditional that if there is unconditional obligation, it must rest in autonomy. In addition, he argues that there really is such an unconditional obligation: "the possibility rests solely on the presupposition of freedom" (V 27:507). But here we reach the end of enquiry. We cannot prove the existence of freedom or explain its exact mechanism. Nonetheless we are justified in assuming freedom because of our awareness of obligation. Kant's reliance on moral obligation to prove freedom, in order then to justify obligation from freedom is not circular: obligation exists if there is freedom, but we know freedom only through obligation. The first relation is a metaphysical grounding, the second an epistemic one (cf. KpV 5:4n.). The smoke lets us recognize that there is a fire, but the fire creates the smoke.

How strong is Kant's argument? It does not need to show that people in fact act freely (cf. KpV 5:45f), but only that we judge an action to be right independently of sensibility. It is therefore no objection if for any actual case in which someone refused to give false testimony (e.g., Thomas More), we can conceive that he might have secretly acted from a desire (for an afterlife, etc.). Kant's point is not to explain an action after the fact. As long as one judges that one ought to act morally independently of one's desires, this Ought would justify assuming Can. Similarly, Kant's point is not

dependent upon any particular example. Any example will do for which one agrees that no desire speaks in favor of the action (the agent does not believe in an afterlife, etc.), and still judges the action to be morally right. One could object that moral views could have been internalized in the course of evolution, and although we judge them to be true, it is not based on a priori reason but our genetic make-up. Kant would reject this possibility since our judgment would not be strictly necessary (cf. KrV B 167f). His point is not whether it feels categorical from the inside, but whether reason would judge the action right independently of our evolutionary history. The crucial question therefore is whether one does judge any action to be unconditionally commanded.

Conclusion

Although obligation is not one of the concepts one would think of first in summarizing Kant, obligation is at the heart of his moral philosophy. He grappled with the problem from the beginning, and autonomy and the categorical imperative are his solution to the problem. Whether he is right will depend on our views regarding an unconditional ought.

The elusive story of Kant's permissive laws

B. Sharon Byrd

In §§ 20–35 of the Vigilantius lecture notes, Kant discusses permissions and obligations. Kant sees all free actions as being obligatory,¹ meaning prescribed or prohibited, or as being permitted (V 27:513.12–18).² Prescriptions create affirmative obligations or obligations to commit certain actions. Prohibitions create negative obligations, or obligations to refrain from committing actions (V 27:511.23–25). Legal duties are generally³ contained in prohibitions and thus are the substance of negative obligations. The primary legal duty is *neminem laede* – harm no one.

What Kant means by “permitted” is not entirely clear at first blush. On the one hand, Kant seems to be referring to morally indifferent actions or *adiaphora*. On the other, Kant speaks of “permitted” actions to mean actions that are generally prohibited but in some exceptional circumstances permitted. For the morally indifferent actions he notes that not all actions are obligatory and a world in which they were would be horribly limiting (V 27:513.25–27). When Kant speaks of such indifferent actions, he gives the example that what I eat, assuming it does not cause indigestion, is not subject to any law limiting my conduct. In this sense an indifferent action is a nonobligatory action. It would seem, at least from the Vigilantius lecture notes, that permissive laws are not needed for this type of action because the action itself is nonobligatory and thus morally indifferent. For the generally

¹ Although the tendency in English might be to read “obligatory” as meaning “required” or “prescribed,” Kant uses it in the same sense as Gottfried Achenwall (see note 7) to mean any act that is either prescribed or prohibited, or any act as to which an obligation is imposed.

² Kant has this same division in the second *Critique* in a table of categories of freedom under the heading “Quality” (KpV 5:66.25–31): “Practical rules of commission (*praeceptivae*), practical rules of omission (*prohibitivae*) and practical rules of exceptions (*exceptivae*).” See too, from *Reflexionen zur Moralphilosophie*: “An action becomes necessary under a prescriptive law (*legem praeceptivam*), impossible under a prohibitive law (*prohibitivam*) and possible under a permissive law (*permissivam*), i.e. an *adiaphoron morale*” (R # 7031, 19:231). All translations in this chapter are my own.

³ Most Kant scholars agree that legal duties are duties of prohibition. As we shall see in [Section 2.2](#), however, some legal duties, at least on the international level, are duties of prescription.

prohibited actions that are exceptionally permitted, Kant speaks of permissive laws (*leges permissivae*). The permissive laws in this case function as justifications for committing otherwise prohibited acts. Several of Kant's examples of this type of permissive law will be discussed in more detail later in this chapter.

The Kant secondary literature has exploded with articles on the permissive law.⁴ Disagreement concerns what the permissive law is, whether there is more than one type of permissive law in Kant's works, whether Kant abandoned one meaning of permissive law in favor of another over time, and when exactly one needs a permissive law at all. This chapter begins with the logic of obligations and permissions (Section 1) and continues with Kant's works and the three permissive laws in Kant's writings (Section 2). It examines Kant's statements in the *Vigilantius* lecture notes (Section 2.1) and expands on that discussion to include all of Kant's examples of permissive laws relating to legally relevant actions (Sections 2.2–2.4). As I will show in this chapter, there are exactly three types of permissive laws and Kant's works contain examples of these, and no more than these, three types. Any other interpretation of a permissive law within the domain of legally relevant actions exceeds the logic of obligations and permissions, a logic with which Kant was familiar.

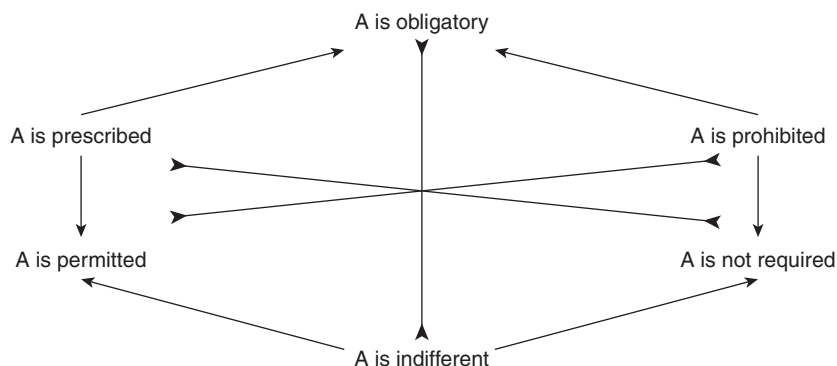
1 The logic of permissions

As Kant notes, the entire division of actions would fail if no actions were indifferent (V 27:513.18f). The division Kant means of prescribed actions, prohibited actions and indifferent actions is the foundation for the deontic hexagon.⁵ The deontic hexagon expresses the logical relations among the propositions: (1) A (an act) is prescribed, (2) A is prohibited, and (3) A is indifferent. From these three propositions, one can derive by implication the other three propositions needed for the hexagon. From the first proposition, "A is prescribed," one can derive the proposition "A is permitted," (in the sense of "A is not prohibited"), since for any action within one system of rules if it is required to commit it, it must also be the case that an actor is permitted to commit it. Otherwise, the system of rules would violate

⁴ Brandt 2004, 69–72; Brandt 1982; Byrd 2010; Byrd and Hruschka 2010, 94–106; Flikschuh 2004; Flikschuh 2000, 113–43; Flikschuh 1999; Hruschka 2004, 46f; Kaufmann 2005; Ripstein 2009, 57–85 and 103f; Roff 2013, Chapter 3; Szymkowiak 2009; Tierney 2001a; Tierney 2001b; and most recently Weinrib 2014.

⁵ Kalinowski 1972, 107f, 119ff examined these six deontic concepts and explained their interrelation.

the principle *impossibilium nulla obligatio est*.⁶ From “A is prescribed” one can also derive the proposition “A is obligatory.” If an act is prescribed within a system of rules, that prescription or requirement imposes an obligation on an actor to commit the action, making the action an obligatory action. From the second proposition, “A is prohibited,” one can derive the proposition “A is not required.” An actor cannot be prohibited from committing an action and simultaneously be required to commit it, again within one system of rules. From “A is prohibited” one can also derive the proposition “A is obligatory,” meaning as to A, actors are placed under an obligation not to commit A. Accordingly, if A is obligatory, then the actor is either required to commit A or required to omit A. The action A is not indifferent. We now have the six corners of the hexagon: (1) A is obligatory, (2) A is prescribed, (3) A is prohibited, (4) A is permitted, (5) A is not required, and (6) A is indifferent. The hexagon can be portrayed as:



If we examine the hexagon, we see that it also contains three pairs of contradictories. A contradictory relationship is expressed when two propositions can neither be true nor false together. “A is prescribed” is the contradictory of “A is not required.” It cannot be the case that “A is prescribed” and “A is not required” are both true because an act cannot be both required and not required by one and the same system of rules. It also cannot be the case that “A is prescribed” and “A is not required” are both false because if

⁶ Assuming that impossible actions are not only those that are physically impossible but also those that are morally impossible because prohibited.

they are both false we have a situation again in which A is both required and not required. In addition, "A is prohibited" is the contradictory of "A is permitted." In one system of rules, A cannot be both prohibited and permitted, nor can it be false that A is prohibited and also false that A is permitted. Finally, "A is obligatory" is the contradictory of "A is indifferent." Either an action is subject to some obligation or it is an indifferent action and no obligation applies to it. An action cannot be both obligatory to commit or omit and at the same time be indifferent as to whether the actor commits or omits it.

As we can see from the hexagon, there are three and only three senses of the word "permitted." An action may be permitted because the action is indifferent, like whether one stands or sits, drinks milk or water. An action may be permitted because the action is prescribed and thus must be permitted to commit, and an action may be permitted in the sense that an actor is not required to commit it because it is prohibited. Although it may seem odd to speak of an action that is prohibited as being permitted, what "permitted" means here is that the actor is permitted to omit or permitted not to commit the action. The action is "not required" but may also be prohibited. Kant's works provide sufficient examples of all three of these meanings of "permitted," all of which will be discussed below.

2 The three permissive laws in Kant's writings

Kant was familiar with the logical relations among these laws from his reading of Achenwall's *Prolegomena Iuris Naturalis*.⁷ Kant's familiarity with these logical relationships is apparent from both the Vigilantius lecture notes and

⁷ Achenwall 1767, § 26, 24f. I am using the text from the third edition rather than the second, which Kant probably used, because the third contains several improvements over the second:

If a free act is seen in reference to a certain class of laws, then it is either the case, or it is not the case, that the act is contrary to one of those laws: in the former case (and as long as the class of laws in question is assumed) it is called an ILLICIT ACT, in the latter case a LICIT ACT. Furthermore, it is either the case that a determination has been made in one of the laws in question that a particular free act should or should not be done, or it is not the case that such a determination has been made: in the former case it is called an OBLIGATORY ACT, in the latter an INDIFFERENT ACT. Finally, an obligatory ACT is either PRESCRIBED or PROHIBITED depending upon whether it has been determined that it should be committed or omitted. Thus on the one hand it is clear: 1) *every non-indifferent act is an obligatory act*, 2) *an indifferent act is licit regardless of whether it is committed or omitted*. On the other hand the following rules result: *with respect to that obligatory act, which is licitly committed, it is illicit to omit it; with respect to that obligatory act, which is licitly omitted, it is illicit to commit it*.

Cf. Hruschka 1986, 43–48; Hruschka 1990. Kant speaks of "erlaubt (*licitum*)" and of "bloß erlaubt (*indifferens*)" in reliance on Achenwall's "actio licita" and on Achenwall's "actio indifferens."

Kant's published works. Still, Kant never spells out the relationships but simply relies on his readers' knowledge of them. Because Kant was familiar with the concepts and their logical relationships, he was also familiar with the three concepts of "permitted." Unsurprisingly, these three meanings of "permitted" in relation to the nature of the act they permit were understood in Kant's time, but few people regard them today. Perhaps for this reason, much disagreement surrounds the nature of permissive laws in the current secondary literature.

Actions that are permitted because they are indifferent are what Kant calls *adiaphora*. Kant also calls this type of indifferent action "merely permitted" (*bloß erlaubt*).⁸ As we shall see later in this chapter, for some types of indifferent actions one does need a permissive law to enable individuals to commit them. For one of the other two concepts of the permitted action, Kant discusses two cases in the *Vigilantius* lecture notes that illustrate a permitted action that is otherwise prohibited. In the first of them, the action is not only permitted in the sense of being justified and thus (exceptionally) not obligatory to omit, but also required, clearly fitting into the category "A is permitted" in the lower left corner of the hexagon. In the second case, the action is permitted as an exception to a primary prohibitory norm, but the action is not required. Nonetheless it too falls into the category "A is permitted" in the lower left corner of the hexagon.⁹ In the *Vigilantius* lecture notes, Kant does not discuss any actions that are permitted in the sense of being not required and which fall in the category in the lower right corner of the hexagon. To find an example of that type of action, one has to look at *Perpetual Peace*.

2.1 *The Vigilantius lecture notes: permissive laws as exceptions to prohibitions*

The two cases discussed in the *Vigilantius* lecture notes are those in which a natural permissive law allows might to precede right, or exceptionally permits the commission of an otherwise prohibited action. One is the case of individuals in a state of nature. In the state of nature, individuals rob themselves of right through reciprocal resistance in protecting what they

⁸ RL 6:223,5–8: "An action that is neither prescribed nor prohibited is merely permitted because with respect to it no law limiting freedom (moral faculty) and thus also no duty exists. Such an action is called morally irrelevant (*indifferens, adiaphoron, res merae facultatis*)."

⁹ The implication relation between "A is prescribed" and "A is permitted" means that all required actions are permitted. Yet not all permitted actions in this sense of "permitted" are also required actions.

judge to be their own. They thus eliminate any possibility of their moving together to a civil or juridical state. This situation is incompatible with the moral law, and in such a case, might precedes right. The strongest among them may force them all to submit to universal law that makes each individual's freedom of choice compatible with everyone else's freedom of choice. In other words, the strongest enjoys the benefit of a permissive law to use might to establish right (V 27:514.25–33),¹⁰ even though the use of force is generally prohibited under the principle *neminem laede*. Later in the *Doctrine of Right*, Kant in fact indicates that not only is the strongest permitted to use might to establish right, but also that he and everyone else have a duty to establish right by moving to a juridical state. This case provides us an example of the sense of “permitted” that is implied by the proposition “A is prescribed.” Everyone is required to move to a juridical state under the postulate of public law (RL 6:307.9–11).¹¹ Indeed, refusing to make this move constitutes an injury to everyone with whom the objector can come into contact. Accordingly, everyone is permitted to use might to coerce everyone else to make that move. Kant sometimes calls this sense of “permitted” *licitum*.¹²

The other case in the *Vigilantius* lecture notes in which might permissibly precedes right is one version of the Plank of Carneades case. In this version, two shipwrecked persons are both trying to obtain possession of a plank that can save only one life. If neither of them can use force to take possession of the plank, then both will die. Accordingly, here too might precedes right in order to establish a right for one to save his life. In other words, the stronger (presumably) enjoys the benefit of a permissive law to use force to gain possession of the plank so that at least one life will be saved

¹⁰ See, too, V 27:515.28–32: “[O]ne must assume that nature in general allows bringing the free choice of human beings in agreement with universal freedom according to a universal law and thus there is here a natural permissive law to use the force applied” and V 27:516.4–7. This situation is what Heather Roff refers to as a “supreme moral emergency” in Roff 2013, Chapter 3, esp. 77–78.

¹¹ One might claim that having a duty to act cannot be a legal duty because it is not the subject of an obligation not to harm others (*neminem laede*). Presumably Kant sees the duty to act to bring about the juridical state in fact to be a negative obligation not to deprive others of security for their rights. Refusing to enter the juridical state places everyone in a state of insecurity and bringing about this insecurity is what is prohibited. Furthermore, Kant plays with the difference between duties of commission and duties of omission: “all acts of commission are acts of omission in a legal sense” (V 27:514.15). Kant also says that the duty to fulfill a contractual obligation is a negative duty even though the obligation requires performance of an act because what is prohibited is depriving the other of what belongs to him, namely performance under the contract: “I cannot pay back a debt without giving [money back to the creditor] but according to the law of freedom one has here only the principle not to take the claim to repayment away from him” (V 27:527.5–8).

¹² RL 6.222.27f: “Permitted is an action (*licitum*) that is not contrary to obligation.”

(V 27:516.20–33).¹³ Again this case falls within the lower left corner of the hexagon. It is generally prohibited under the principle *neminem laede* to use coercion against one shipwrecked person to prevent him from gaining possession of a plank. Yet, in the exceptional case in which both will die if one does not use coercion to secure his own possession of the plank, then that person has a permissive law in the sense of a justification to use coercion to get the plank into his possession. This act, unlike the act of founding a juridical state, is not required. Nonetheless the act falls within the lower left corner of the hexagon because it is permitted in the sense of being the contradictory of being prohibited. Accordingly, although all required actions are permitted in the sense of the lower left corner of the hexagon, not all permitted actions are also required actions. Some permitted actions are permitted simply in the sense of being not prohibited because they cannot be prohibited in the exceptional circumstances.¹⁴

From these two cases Kant concludes that there are natural permissive laws.¹⁵ The rule for deciding whether a natural permissive law allows might to precede right is whether right can be brought about without might preceding it. If not, a natural permissive law allows might to precede right (V 27:515.5–10). In the case of establishing a juridical state, Kant assumes that individuals will not voluntarily give up their savage freedom and submit to coercive public laws. If not, then right, which can be ensured only in a juridical state, cannot be attained and a natural permissive law allows using coercion to establish the juridical state. In the case of taking the disputed plank to save one's own life, might also precedes right. Without a permissive law, neither of the two persons shipwrecked would have a right to life unless one of them can establish his right to the plank by successfully taking control of it. Once he has done that, then he does have a right to possess and defend the plank against the other's attack.

Kant also points out that positive or statutory law can include permissive laws. That is because positive law contains general prohibitions that admit of exception, in contrast to universal prohibitions that do not. In such a case

¹³ In contrast, Kant indicates that if one person is already on the plank the other cannot push him off in order to save his own life. If he could, then he would have a coercive right against the plank possessor and the plank possessor would have a coercive right against the attacker to protect his possession of the plank and these two coercive rights would be contradictory (V 27:516.12–20).

¹⁴ It is highly doubtful that Kant maintained his position on this case in 1797. In the *Doctrine of Right*, he deplores the utilitarian calculus that permits one person to be sacrificed even to save an entire people (RL 6:331.31–332.3). See, too, [Section 2.4](#) and the arguments presented there on the utilitarianism involved in the plank cases.

¹⁵ It is this concept of “permissive law” on which Reinhard Brandt focuses in Brandt 1982 and Brandt 2004. Katrin Flikschuh's approach is similar in Flikschuh 2000, Flikschuh 1999, and Flikschuh 2004.

the permissive law also functions as a justification or an exception to a norm prohibiting certain conduct (V 27:514.17–21). In addition, Kant indicates that permissive laws can relate to indifferent actions or *adiaphora*. *Adiaphora* are actions that are not the object of obligation, for example, walking in my yard or sitting still. They have no relation to a moral law that determines my use of freedom. Indeed, reason leaves the action completely up to the actor's free choice and thus no obligation hinders the actor in committing the action (V 27:512.11–24). Kant claims that one must assume there are permissive laws because otherwise all actions would be subject to prohibitions or prescriptions and no action would be indifferent (V 27:513.15–30).¹⁶

2.2 'Perpetual peace': permissive laws as exceptions to prescriptions

In *Perpetual Peace*, Kant discusses three cases in which an action is required but because of the exceptional circumstances can be (temporarily) omitted justifiably. These cases fall within the lower right corner of the hexagon because they concern justifications for not performing a required act. The cases are contained in Preliminary Articles 2, 3, and 4. Preliminary Article 2, for example, provides that "No existing independent state . . . shall be acquirable by another state through inheritance, exchange, purchase or donation." The prohibition against this acquisition is "universal," as Kant calls it in the *Vigilantius* lecture notes (V 27:514.17–19), because it admits of no exception. Nonetheless, in *Perpetual Peace* Kant states that the requirement to restore the *status quo ante* the acquisition does include a permissive law that allows gradual rather than sudden restoration of the acquired state's independence.

In a most recent article, Jacob Weinrib argues that this permissive law is different from the two Kant discusses in the *Vigilantius* lecture notes because here the permissive law is not a justification (see Weinrib 2014). Weinrib relies heavily on Kant's statement with regard to these preliminary articles that they do not contain "exceptions to the legal rule" (ZeF 8:347.20f). It is certainly true that a state after adopting¹⁷ the preliminary articles in *Perpetual Peace* cannot under any circumstances acquire another state in one of the ways listed. In this sense I agree with Weinrib that the legal rule does not admit of exception. Nonetheless the permissive law that allows a state that

¹⁶ Kant seems to change his position on whether one even needs permissive laws for indifferent actions in the RL 6:223.9–17. For a discussion of this passage in the RL, see Hruschka 2004.

¹⁷ *Perpetual Peace* in its preliminary and definitive articles is written like a treaty for states to ratify if they are aiming at attaining perpetual peace.

already has acquired another to only gradually restore the acquired state's independence is an exception to the rule that independence must be restored. The exception says that although generally independence must be restored, still in cases in which sudden restoration would bring more evil than good, the restoration may proceed gradually. As Kant himself says, "the prohibition relates only to the future mode of acquisition . . . the release from this prohibition, i.e. the permission, relates to the current status of possession" (ZeF 8:347.29–33). Accordingly, the prohibition in Preliminary Article 2 is absolute, or universal, and does not permit of exception. This prohibition relates to the future "mode of acquisition." Yet we also have another rule relating to the "status of possession" requiring restoration of independence. This rule does admit of exception. The rule is in the form "A is prescribed" – A being restoration of the acquired state's independence – and the permissive law permits not committing (at least for a while) an otherwise required action. Therefore, we have an example here of a permissive law permitting an action in the sense of the lower right corner of the hexagon.¹⁸ Here, I also agree with Weinrib that the permissive law in *Perpetual Peace* is different from the permissive law in the *Vigilantius* lecture notes. Still the difference lies not in the fact that the permissive law in *Perpetual Peace* is not a justification. The difference lies in the fact that the permissive law in *Perpetual Peace* justifies making an exception to a *prescription*, rather than an exception to a prohibition as in the two examples discussed in the *Vigilantius* lecture notes.

2.3 'The doctrine of right': permissive laws as power-conferring norms

Joachim Hruschka can be credited with first discovering the third type of permissive law (see Hruschka 2004). This permissive law relates to indifferent actions that are accordingly neither prescribed nor prohibited. They are thus actions that are not obligatory. Still, as Hruschka argues, committing one of these actions leads to the creation of rights and thus to the imposition of legal obligations on others not to violate those rights. This type of permissive law is found in § 2 of the *Doctrine of Right*: "It is possible to have any external object of my choice whatsoever as mine, i.e. a maxim, which if it became law, meant that an object of choice must become *per se*

¹⁸ Brandt does not consider the difference between a justification for committing a prohibited action and a justification for omitting a required action, Brandt 1982 and Brandt 2004.

(objectively) *masterless* (*res nullius*), is wrongful” (RL 6:246.5–8). Hruschka interprets this permissive law as being a power-conferring norm that gives us an authorization or moral faculty to acquire and exclusively possess external objects of choice, namely physical things external to us, claims under a contract, or claims to determine someone else’s choice to commit an action, and another person’s status in relation to me, such as a person’s status as *my* spouse or a young person’s status as *my* child. Through this permissive law we can become the owners of external things, claimants under a contract, spouses or parents. According to Hruschka, the permissive law in the *Doctrine of Right* expands our freedom of choice beyond our innate right to freedom from external interference with our physical integrity to include external freedom in the exclusive use of objects of choice. Although acquiring an external object of choice as mine is an *indifferent* action, meaning I am neither subject to a prescription to acquire external objects of my choice nor subject to a prohibition to not acquire them, still once I have acquired an external object and claimed it as mine, others have an obligation to refrain from taking and using that object.

Some authors¹⁹ conclude that here too the permissive law is a justification just as it is in the Vigilantius lecture notes and in *Perpetual Peace*. It justifies my unilaterally taking an external object of choice as mine and imposing an obligation on others not to disturb my exclusive possession of it. The justification is argued to be necessary because the initial taking violates others’ freedom (of choice) by placing a usable object beyond their possibility of use. The obligation the acquirer then imposes on others not to disturb his possession of the acquired external object results from the wrongful act of acquisition.²⁰ Yet this approach fails to give proper attention to the meaning of choice for Kant and to the postulate itself.

¹⁹ Among others Flikschuh 2000; Roff 2013.

²⁰ See, e.g., Roff 2013, 66–67: “Now, we not only have the problem of . . . unilateral acquisition violating freedom, but also an *obligation* imposed on others by the (wrongful) act of unauthorized unilateral acquisition. . . . The postulate as *lex permissiva*, therefore, allows a violation of the Universal Principle of Right to establish the conditions for right (as it relates to the exercise of freedom).” One might point out that Roff’s quotation of the Universal (or general) Principle of Right leaves out any reference to the word “choice”: “any action is right if it can coexist with everyone’s freedom in accordance with a general law” (157, note 24 to Chapter 3). This omission is perhaps not so fatal, and it corresponds in part to the Mary Gregor translation, because as Kant states: “Only choice can be called free,” MS 6:226.10f. Hence, “everyone’s freedom” refers to everyone’s freedom of choice. See also Flikschuh 2000, 141 relating a scenario of an original acquisition in which a nonacquirer appeals to the argument that the acquirer has violated the non-acquirer’s “equally valid right” to have the external object of choice the acquirer has taken for himself and not to accept an obligation to refrain from using that object. According to Kant, the non-acquirer does not have any right to the external object but merely a possible wish to have it.

Kant contrasts choice to wish. A person has a choice with respect to an object of his desire only if he is aware that his self-determined action can bring forth that object. He has a *choice* to have an apple if he sees the apple and knows he can get to it and take it before anyone else gets there to interfere. In contrast, if he cannot bring forth the object of his desire he has only a *wish* to have this object.²¹ An actor may wish to have an apple but know the apple is closer to me and I am about to take it. If I do take it, I have not violated his freedom of choice, because with respect to taking the apple he had no choice, but only a wish. If I do not violate his freedom of choice I need no justification for taking the apple. Consequently, the permissive law in § 2 of the *Doctrine of Right* cannot be a justification for committing a prohibited act as the permissive law is in the Vigilantius lecture notes or for omitting a required act as it is in *Perpetual Peace*. Furthermore, the postulate states that a law according to which an external object of choice must become *per se* masterless is wrongful. Thus “[o]ne who proceeds following a maxim through which it would be impossible to have an object of my choice as mine injures me” (RL 6:256.25–27).²² By acting wrongfully that person violates the acquirer’s right to external freedom of choice and thus gives the *acquirer* a justification to use coercion to enforce the obligation under Kant’s idea that a right is necessarily coupled with a right to coerce (RL 6:231.32–34). This justification, however, is not a justification for violating others’ freedom of choice by unilaterally acquiring an external object because according to the postulate, that act is not wrongful. Instead it is a justification for using force to enforce one’s own rights to the external object one has rightfully acquired, or the justification of self-defense.

Another line of criticism of Hruschka’s work accepts his theory that the permissive law in the *Doctrine of Right* is a power-conferring norm but disagrees that Kant abandoned his position in the Vigilantius lecture notes and *Perpetual Peace*.²³ This criticism appears to be correct since indeed three different kinds of permissive laws logically exist as can be seen in the deontic hexagon, with whose concepts Kant was familiar, and also, as this chapter has illustrated, in Kant’s works.

²¹ MS 6:213.17f: “To the extent it [a faculty to do or leave undone as desired] is connected with the awareness of the capacity of one’s act to bring forth the object, it is called *choice*; if it is not so connected then the act is called a *wish*.”

²² This same duty is formulated also in § 6, RL 6:252.11–15, where Kant says that the duty follows directly from the postulate of practical reason.

²³ Most notably, Weinrib 2014.

2.4 The preliminary work on the *Metaphysics of Morals*: permissive laws

For the sake of completeness, three additional examples of permissive laws in Kant's preliminary work on the *Metaphysics of Morals* (VMS 23:385.25–27) should be mentioned. Kant gives no explanation of his reasons for selecting these three cases or any analysis of why permissive laws allow the commission of these otherwise prohibited acts. Kant simply says that “the permissive law is a law through which something is permitted according to natural laws that is prohibited by civil laws.” The cases are: (1) being one's own judge in cases of injury, (2) polygamy when one is the last man on earth, and (3) robbery if one is in danger of starving. The latter two examples of permissive laws are easily located in the lower left corner of the deontic hexagon (*licitum*). They are thus no different from the permissive laws in the Vigilantius lecture notes. Both actions, polygamy and robbing someone else, are generally prohibited but exceptionally permitted because of the extreme situation.

For case two, the last man is permitted to engage in polygamy contrary to positive law, presumably in order to preserve human beings from certain extinction on earth. Here positive civil laws would prohibit polygamy, but a natural law allowing protection of the existence of the human race would exceptionally permit it. Case three also is exceptionally permitted, at least according to Kant in his preliminary work on the *Metaphysics of Morals*, to allow a person in danger of starving to save himself. This case, however, seems clearly at odds with Kant's later position in the *Doctrine of Right*. In the *Doctrine of Right*, Kant rejects any utilitarian calculus from his legal philosophy. One example he discusses is again the Plank of Carneades case, but in contrast to the case discussed in the Vigilantius lecture notes, in the *Doctrine of Right*, the case involves one person who has already gained control of the last remaining plank being pushed off the plank by another to permit the other to save his own life. Kant points out that the person holding onto the plank has not been an aggressor against the other and thus the other cannot claim any right to self-defense in pushing the holder of the plank off of it. Although Kant indicates that punishment would not be appropriate in such a case,²⁴ Kant's solution is to give the aggressor an excuse but not a justification for his action. He will not be punished but

²⁴ The reason why punishment is not appropriate is that the criminal law's threat of punishment could not possibly have deterred the aggressor from acting. In order to understand why that is, see Byrd 1989 and Hruschka 1991.

nonetheless his act was not permitted. Accordingly, Kant rejects the existence of a permissive law in this type of case to allow someone to commit a wrongful act.

One might object to a comparison of this Plank of Carneades case to the third case in the preliminary work because here the aggressor can save his own life by sacrificing a mere property interest of the person whose food (or money to buy food) he takes. The utilitarian calculus would permit sacrificing property to save life but not life to save another person's life, as in the plank case. Still, in the *Doctrine of Right*, Kant also rejects a utilitarian calculus to justify using a condemned volunteer prisoner for dangerous medical experiments in return for preserving his life by not executing the death penalty imposed. Kant indicates that the interests being balanced here are the medical community's opportunity to get an insight, which could benefit countless other individuals, as opposed to justice and particularly punishment justice. The most likely conclusion one can draw from Kant's rejection of utilitarianism in the *Doctrine of Right* is that at least as of 1797 he would no longer consider a permissive law appropriate for the latter two cases he discusses in the preliminary work. Nonetheless the cases can be classified in terms of the categories in the deontic hexagon and do not represent some fourth type of permissive law.

The first case in the preliminary work is more puzzling. Generally Kant rejects acting as judge in one's own case. One of the major defects of the state of nature is that there is no public judge who can decide a dispute between two individuals, both of whom may be acting in good faith according to their own lights of what is good and just. Since each of the disputants may be influenced by his own desires, the resulting judgments may differ and the only enforcement mechanism would be brute force. Such a state is a state of war where no one is given any security for his rights (RGV 6:97.31–34). In the juridical state, or civil condition, a neutral judge is available and acting as judge in one's own case would be prohibited by positive laws as a form of vigilante justice. Nonetheless, a permissive law could allow one to judge in one's own case, if no judge is available. A situation is easily imaginable in a juridical state. If A attacks B in a dark alley where no police protection is to be had, then B will have to act as judge in his own case, namely he will have to determine that A is violating his right to physical integrity, before he can exercise self-defense against A. One may view such a situation as a temporary state of nature within the juridical state because the institutions that characterize the juridical state are momentarily lacking. If B, or people like him, cannot judge in his own case then generally people whose rights are attacked

would be left helpless. To avoid injustice because of the unusual situation, a natural permissive law would permit what otherwise would be prohibited by positive law. Here again we have an example of a permissive law that falls in the lower left corner of the deontic hexagon (*licitum*). The permissive law permits what otherwise is prohibited.²⁵

3 Conclusion

Kant used three and only three concepts of a permissive law in all of his writings. He begins in the *Vigilantius* lecture notes with the permissive law as a justification or exception to a prohibitory norm, or the permissive law in the sense of *licitum* (lower left corner of the hexagon). He continues in *Perpetual Peace* with the permissive law as a justification or exception to a prescriptive norm (lower right corner of the hexagon). Finally, in the *Doctrine of Right*, he discusses the permissive law as a power-conferring norm, which expands our moral faculties to include the faculty to have external objects of our choice as our own. This permissive law applies to otherwise indifferent actions (bottom center corner of the hexagon) but through acting under the authority of this permissive law, the acquirer creates legal obligations for all others. There are no other categories of permissive laws as can easily be seen in the logic of the deontic hexagon. Kant has employed them all in his legal writings.

²⁵ Kant discusses two more permissive laws in the *Doctrine of Virtue*. The first is a permissive law that allows using a marriage partner for pure pleasure even though propagation – the natural end or purpose of sexual intercourse – is impossible because of pregnancy, age or illness (TL 6:426.7–15). The second is a permissive law that allows a person to not render aid to someone in distress (TL 6:453.6–15), but Kant rejects this latter permissive law as self-contradictory. Because the first permissive law is within the realm of ethics rather than law, it is beyond the scope of this chapter. To accommodate ethical laws, one needs the deontic decagon. See Hruschka and Joerden 1987.

On the logic of imputation in the Vigilantius lecture notes¹

Joachim Hruschka

In his lecture of the winter semester 1793–1794, Kant discusses the logic of imputation. Imputation became a focal point in moral philosophy when Samuel Pufendorf first theoretically analyzed it in his *De Jure Naturae et Gentium* of 1672.² In the aftermath of Pufendorf's work, "imputation" came to mean a judgment that a subject reaches regarding his own or a cosubject's actions.³ This concept of imputation thus diverts attention away from the (imputed) action to the imputing observer's judgment of this action. For practical philosophy, this diversion of attention from the action to the judgment of imputation corresponds to the (later) shift in attention from the things themselves to the cognizing subject's cognition of these things, namely Kant's famous Copernican Revolution for theoretical philosophy.

As in his earlier lectures, Kant begins with distinctions and divisions that we can find in Achenwall's and Baumgarten's works. Kant used their texts as a basis for his own lectures,⁴ and indeed the Latin terminology in the Vigilantius lecture notes originates primarily from their work. Both of these authors, but more so Baumgarten, distinguish between imputation of an act (*imputatio facti*) and imputation of the law (*imputatio legis*). The imputation of an act is the judgment through which someone is seen as the author of a certain act. The imputation of the law is the application of the law

¹ This chapter was translated from German by B. Sharon Byrd, who also provided the translations of all quoted passages.

² Pufendorf 1989 [1672], 101–5 (Lib. I Cap. 9).

³ Pufendorf speaks of "declarare." Christian Wolff, in *Philosophia practica universalis I*, § 642, uses "iudicare" (Wolff 1971 [1738], 430). In his lecture, Kant too speaks of "*judicium imputatorium*" and of "judgment (*Urtheil*)" (V 27:559.37, 564.1), and he otherwise assumes that imputation is a judgment (cf. MS 6:227.21). When the Vigilantius lecture notes refer to a certain "character of the action" (!) as "*imputatio facti*" (V 27:562.3f), then this designation is obviously false and can be the result only of a mistake in the notes.

⁴ Baumgarten 1760 (here cited from AA 19:7–91); Achenwall 1763.

to this act, or the subsumption of the act under the law.⁵ I shall return to discuss the details of both of these types of imputation later in this chapter.

In his lecture, Kant refers to the distinction between *imputatio facti* and *imputatio legis*. According to Kant, imputation concerns the concept of an *auctor circa factum* (author in relation to an act), on the one hand, and the concept of *meritum* (merit) in the broader sense of “merit” (V 27:558.27f),⁶ on the other, and therefore includes merit (*meritum* in the narrower sense) and demerit (*demeritum*) (V 27:558.34). Decisive for the judgment of merit or demerit is what the law requires the individual to do. If his actions exceed what the law requires, then they will be imputed to him as merit. If they fall short of what the law requires, they will be imputed to him as demerit (V 27:558.27–36; 561.19–21). It follows that the “observance of a legal duty . . . is connected to no merit” (V 27:560.11–14). Although someone’s action will be imputed to him (as his act) when he acts in accordance with the law, the action will not be imputed to him as merit or demerit.⁷ This evaluation is traditional. In 1660 Pufendorf writes “No merit arises from the commission of an act that is owed.”⁸

The text from Kant’s lecture on the observance of a legal duty is predicated on the distinction between “legal” or “owed duties” (V 27:548.4), and “ethical” or “meritorious duties,” which are also called “duties of beneficence” (V 27:548.4 and V 27:560.8f, 560.35f) and with which we are familiar from the *Groundwork of the Metaphysics of Morals* (G 4:424.10–14). Legal duties correspond to another person’s rights and are thus connected with a “coercive right,” or a right to coerce performance of the legal duty (V 27:548.4–7). Their fulfillment is “absolutely mandatory” (V 27:560.15). The fulfillment of ethical duties is “not absolutely necessary” and it “remains within the actor’s discretion how and in what way they should be fulfilled” (V 27:548.4–9). Consequently, the violation of a legal duty is imputed to the actor’s demerit, whilst the fulfillment of an ethical duty is imputed to the actor’s merit.

This chapter first considers the concept of a *causa libera*, which plays a significant role in distinguishing between imputable and nonimputable actions (Section 1). In Section 2, I deal with the relation between *imputatio*

⁵ As Baumgarten 1760 states in § 125 (19:61.22–26): “1) Iudicium, quo quis certi facto auctor iudicatur esse. 2) Applicatio legis ad factum, s[ive] facti sub lege subsumptio.” Cf. also Achenwall 1763, § 29.

⁶ The broad concept of *meritum* is the genus for “meritum,” (merit) in the narrower sense, on the one hand, and for “demeritum” (demerit), on the other. On the broad concept, see Achenwall 1763, § 27.

⁷ The debtor who pays back his debt on time acquires no merit (V 27:560.20–22). The general who orders enemy soldiers to be killed is not blameworthy (Kaehler 88.10–18).

⁸ “Ex actione debita nullum est meritum” in Pufendorf 1999 [1660] 111 (Def. XIX, §§ 1, 2; Def. XX, § 1).

facti and *imputatio legis*, and then proceed to the reasons why imputation could be ruled out (Section 3). Finally, in Section 4, I shall answer the question of why Kant no longer speaks of the *imputatio legis* in the *Metaphysics of Morals*.

1 The concept of a *causa libera*

In the Vigilantius lecture, Kant shows particular interest in the determination that no merit arises on the fulfillment of a legal duty, a determination that he repeats in the *Metaphysics of Morals* (MS 6:227.30–228.2).⁹ We can find Kant's reason for that in the lecture itself. It is not up to the actor, Kant says, to fulfill the duty. Instead the actor "is subjected to the duty and is thus not *causa libera* (author) of his act" (V 27:560.18–22). Accordingly, a person who fulfills a legal duty is not "*causa libera* (author)" of his act, for which reason he acquires no merit. As will become apparent, this claim is problematic.

Kant first discusses the *causa libera* in his lecture shortly preceding the text just quoted. There he says: "Imputation consists . . . of the *judicium aliquem esse auctorem alicujus vel boni vel mali*. It takes the law into consideration, which provides whether something is good or bad, and contains the judgment that the actor is *causa libera vel auctor actionis (facti)*" (V 27:559.36–38).¹⁰ "*Causa libera*" is a technical term, which Christian Wolff introduced into the theory of imputation.¹¹ Achenwall states: "When someone experiences a good or bad result from a certain free act and he could foresee physically that this result would occur, then the act as well as its result is to be imputed to his free choice and the actor is called the 'free cause' or 'author' of both the act and its consequence."¹² Hence the concept of the free cause (in the Wolff/Achenwall tradition) more precisely specifies the concept of the author. A person is the author of an act and author of the consequences of this act when he can be seen as the free cause

⁹ The determination is expressly formulated at MS 6:390.30f: "The commensurability of the action to the law (to be a juridical human being) [is] not meritorious."

¹⁰ "*Judicium aliquem esse auctorem alicujus vel boni vel mali*" means the "judgment that someone is the author of something good or evil." "*Causa libera vel auctor actionis (facti)*" means "free cause or author of an action (deed)."

¹¹ Wolff 1971 [1738] § 526, § 527. The concept, however, is not generally used in the theories of imputation in the eighteenth century. Baumgarten and Daries, for example, do not use it.

¹² Achenwall 1763, § 27: "Si cui ex certo facto libero nascitur consecrarium bonum vel malum, quod sibi ex eo exstiturum praevidere physice potuit: libero eius arbitrio tribuendum est et factum tale et eiusdem consecrarium, indeque et agens vocatur et facti et consecrarii talis CAUSSA LIBERA seu AUCTOR."

of the act and its consequences. Kant proceeds from these concepts both in his lecture and in his later *Metaphysics of Morals*. Kant's reliance on these concepts can be seen at V 27:559.36–38, but also in the following quotation: "Causality in committing the act makes the act a *factum*, and *causa facti libera qua talis* . . . determines the *auctorem facti*" (V 27:561.35–37).¹³ This statement corresponds to Kant's calling the "author" the "*causa libera*" (using the Latin expression) in the *Metaphysics of Morals* and saying that in the case of such authorship the act is called a "deed" (*Tat*) (MS 6:227.21–23).

In line with the tradition of the time, Kant thus distinguishes between "action" (*actio*, *Handlung*) and "deed" (*factum*, *Tat*). An action is an event occurring in the world in which generally, but not always, a person is involved. An action, however, is a deed only if it is the effect of a "*causa libera qua talis*" (the effect of a free cause as such), i.e. only if it is "chosen by free will in accord with a law of freedom" (V 27:561.29–37).

Decisive for imputation is whether the actor "chose" the action "with free will." "*Causa libera*" is contrasted with "*causa causata*" (the caused cause) (cf. R #6019, 18:425.14f). If a human action which is to be imputed is a free action, then the action can be imputed (as a "deed" – *factum*). If the human action is a caused action then it cannot be imputed.¹⁴ A person "[a]s author begins a series of actions, whereby the beginning and the cause lay in himself, not in nature" (V 27:559.8f).

The question arises of when an action is free and when it is not. What Kant has to say in the lecture on (internal) freedom is exceptionally laconic. Imputation, as Kant says there, "considers the law which determines whether something is good or bad." Here "law" means the "laws of freedom from which duties arise, which the individual can fulfill or fail to fulfill" (V 27:559.26–38). The laws are called "laws of freedom" because these laws are contrasted to laws of nature. Actions can be imputed only when we consider them under laws of freedom. Accordingly, the concept of freedom that Kant uses in the lecture is the same as he uses in the *Groundwork of the Metaphysics of Morals* (G 4:446.5–447.25) and also in the *Metaphysics of Morals* (MS 6:213.35–214.4). One must distinguish (according to the *Groundwork* and the *Metaphysics of Morals*) between a negative and a positive aspect of freedom. The negative aspect is the independence of the action "from nature" (cf. V 27:559.30f). An action's independence from nature is only possible when the action is subjected to

¹³ The sentence says that "the free cause of a deed as such determines the author of the deed."

¹⁴ Or as Kant says: "the effect of a *causa naturalis qua talis* (following from natural causes)" (V 27:561.32f).

laws of freedom. This idea of the action being subjected to laws of freedom is the positive aspect of freedom. Laws of freedom originate in reason (cf. V 27:557.38f). Consequently they prohibit and require categorically, that is, one considers only, as Kant states in the lecture, the “moral necessitation through the law of freedom,” and not (natural scientific) necessities in the empirical world, including forces that originate in hypothetical imperatives.

The “*causa libera*” of an action is thus a person in light of his duties, that is, whether in a concrete situation that person fulfills or fails to fulfill his duties (cf. V 27:557.31–558.3). These duties include legal duties. Thus the person who fulfills his legal duties is the *causa libera* (free cause) of the fulfillment of these duties and author (*Urheber*) of the action through which he fulfills these duties. This statement, however, at least in the first conceptual move, is contradictory to the thesis in V 27:560.18–22 that one who fulfills his legal duty is *not* the *causa libera* and author of his actions.

Of course, one could say that someone who fulfills his legal duties is not free in a particular respect *because* he is subject to a legal duty. Under this assumption one could also say that the actor is not the *causa libera* of the legally rightful action. Yet then the word “*liberum*” in the expression “*causa libera*” in V 27:560.18–22 has a different meaning from the “*liberum*” in the expression “*causa libera*” in other places in the lecture. At V 27:560.18–20 the actor is free only if he was not subjected to any legal duty. In other places in the lecture, however, an actor is free if he is subjected to a duty, including a legal duty. As other lecture notes (by Kaehler, Mrongovius, and Collins) on the topic of freedom show, Kant realized that two different meanings of “*liberum*” were conceivable, but the Vigilantius lecture notes do not reveal Kant’s position on this issue.¹⁵

If an action was a legal duty, that is, “not free but coerced by the law,” then the action cannot be imputed to the actor (Kaehler 89.25–90.4; M I 27:1438.6–8; C 27:289.13f). In the other lecture notes, however, the authors clarify by saying that in such case the *imputatio facti* is not excluded, but instead the *imputatio legis* (see [Section 2](#) on this distinction). If one in this case (namely when the *imputatio facti* is not excluded but the *imputatio legis* is) wants to say the actor was not the *causa libera* of his action, then that

¹⁵ Of course one could think that the passage at V 27:560.18–22 is the result of a mistake in transcription. Yet the ambiguity of the expression “*liberum* = free” repeats itself in the corresponding ambiguity of the word “*posse* = can” in the statement “*ultra posse nemo obligatur*” (see [section 3](#)). In light of the other lecture notes cited here it seems clear that Kant himself was aware of the ambiguity. Consequently, one must assume that Kant discussed the problem but his comments on it did not make it into the lecture notes.

is indeed a conceivable position. Of course we are using two different concepts of freedom here. Correspondingly, one could also speak of two different concepts of author. Author is one to whom an action can be imputed as a deed (*factum*). If someone is an author of an action in this sense, then this judgment is retracted (the person is thus *not* the author) if the *imputatio legis* is excluded. Still, this way of distinguishing between two meanings of “author” seems counterintuitive.

In the lecture, Kant seems to be playing with the concept of the *causa libera*, to which he refers eight times (V 27:559.36–38, 560.19, 561.12, 561.23, 561.33f, 561.36, 563.34, 565.29). The concept of the *causa libera* is mentioned only once in the *Metaphysics of Morals*, namely when Kant in his definition of imputation uses it to explain the concept of the author (MS 6:227.21–23). Freedom is a condition for the possibility of duty, and in the *Metaphysics of Morals* Kant, to avoid any misunderstanding, obviously no longer wants to say that a person is not free when he fulfills his duties.

2 The difference between *imputatio facti* and *imputio legis*

The determination that an action is a deed, that is, the *imputatio facti*, does not indicate that an *imputatio legis* has occurred, or that the potentially applicable laws have in fact been applied to the deed. If a law is applied to a deed then we are confronted, as Kant emphasizes in the older lecture notes, not with one but, rather, with *two* imputations (Kaehler 88.7f; C 27:289.5f; M I 27:1438.1f). In the *Vigilantius* lecture notes, Kant discusses these two levels of imputation extensively, particularly the *imputatio legis*. Imputation of the law presumes “that the action is subject to a law . . . and can be subsumed under this law.” It is the “*applicatio legis ad factum sub lege sumtum*,” that is, the application of the law to the deed that has been subsumed under the law. The imputation of the law can thus be characterized as the conclusion of a practical syllogism. Baumgarten speaks of a “syllogismus imputatorius.” The law provides the major premise; the deed (*factum*) provides the minor premise (Baumgarten 1760, § 171 [AA 19:79.4–7]). The conclusion is, as we can add from the *Critique of Pure Reason*, “the real judgment, which expresses the assertion of the rule [i.e. the assertion of the law] in the subsumed case” (KrV 3:255.29f = A 330/B 386f). Kant provides an example: “The defamer should return the victim’s honor. – He defamed me. – Therefore he must provide satisfaction” (V 27:562.5–18). The logic behind this conclusion is straightforward. The intellectually interesting part is hidden in the minor premise, which requires judgment to construct.

These two imputations are not so easy to distinguish, as will be seen. Kant, as Baumgarten, pays particular attention to certain details of the imputation of the deed (*imputatio facti*). In this connection, the key concept is the “*species facti*.” The *species facti* is a commonly used technical term, which was used in the eighteenth century and earlier in court decisions as the title of explications of the facts determined (cf. Schott 1965, 174). In the eighteenth century the *species facti* was the focus of theoretical discussions. Authors such as Daries, Baumgarten, and Martini wrote on this topic.¹⁶ In his lectures, Kant is one of the last in this development. He uses the expression as if its meaning could be assumed to be self-evident.¹⁷

“*Species facti*” does not mean, as one might assume *prima vista*, “species of deed” (as opposed to the complementary genus), but, rather, “picture of the deed.” Daries also uses concepts such as “*imago*” and “*effigies*” in this connection (Daries 1754, Obs. XLVI, § 1). Daries defines the *species facti* as a “list of characteristics that constitute the deed under discussion as such.”¹⁸ Baumgarten defines it as a “list of the essential moments of a deed,” “*enumeratio momentorum in facto*” (Baumgarten 1760, § 128 [AA 19:62.26]), a formulation that Kant also uses (V 27:563.9f).

We must understand the *species facti* as the answer to a question, the *quaestio facti*. Baumgarten sees it in this way (Baumgarten 1760 § 128 [AA 19:62.24]), as does Kant after him (V 27:562.24f). The *quaestio facti* is the question about the (legally) relevant circumstances surrounding an action, which Baumgarten calls “*momenta in facto*” (Baumgarten 1760, § 128 [AA 19:62.23–26]; Kant: V 27:562.34f). The answer to the question about the essential moments of a deed is the *species facti*. The “*essentialia facti*,” “the essential aspects of the deed” and the “*momenta in facto*” are, as Kant says, one and the same, whereby the “*extraessentialia*” (the unessential aspects of the deed) have nothing to do with the *quaestio facti* itself and are not part of the *momenta in facto* (V 27:563.3–8).

We now have the decisive question of from where the essential nature of the facts comes. The answer that Kant gives is: “When determining the *circumstantias in facto* it is necessary to consider the law in order to find the *momenta in facto* since even though here the law is not yet imputed it contributes to a more complete determination of the *facti* itself”

¹⁶ Cf. Daries 1740, 129 (§ 212); Daries 1754, 130–3 (Obs. XLVI); Baumgarten 1760, § 128 (AA 19:62.19–34); and Martini 1770, 148 (§ 174). On the details, see Hruschka 2001.

¹⁷ In TL 6:425.13 it says that human beings have no qualms about revealing a suicide “with all of its atrocities (in a *species facti*) to the whole world.”

¹⁸ “*Enumeratio characterum, qui rem obviam, qua tale, constituunt*,” Daries 1754, Obs. XLVI, § 2.

(V 27:563.14–17).¹⁹ Similarly a few pages later on the “*factum*,” the question arises “whether the law is the one under which the *factum* can be located and [whether] the *factum* has such elements that it can be subordinated to the law” (V 27:572.14–19). In other words: the relevance of a deed and its circumstances follow from the law *before* any application of the law. With this comment Kant, who far surpassed Baumgarten, has hit upon what in today’s jurisprudence is referred to as the “hermeneutic circle.” In the twentieth century, Karl Engisch described this circle as the “casting back and forth of one’s sight between the major premise and the facts of a case” (Engisch 1960, 15), whereby the major premise, as for Kant, is the law and Engisch designates the deed, together with its relevant circumstances, as the “facts of a case.” To correctly choose the facts and to interpret them, the person sitting in judgment must have a look at the law *before* applying it, and to find the relevant law the judge must have a look at the facts that are to be subsumed under the law, again *before* applying the law to them. Kant limits himself to the first aspect of the circle. He does not speak in the same clarity of the selection and interpretation of the law in light of the facts, although we can assume he also saw this aspect of the problem.

The lecture notes at V 27:572.14–19 are under the title “*lex et factum*” (“law and deed”). The title could also designate the text at V 27:563.14–17. Kant retains his interest in the topic of the relationship between the law and the deed into the *Metaphysics of Morals*. Here he distinguishes between the “*lex iusti*” and the “*lex iuridica*.” The *lex iusti* is the law of what is right according to external laws, or shorter, the law of what is right, be it the applicable natural law in a certain situation, be it the positive law to be applied in this situation. The *lex iuridica*, in contrast, is the legal nature of a concrete situation; it is reality (*Wirklichkeit*) that is seen from a legal viewpoint. In the *Metaphysics of Morals*, Kant specifies the relation between the *lex iusti* and the *lex iuridica* as the relation between form and matter (MS 6:306.8–11). Reality is legally stamped by the *lex iusti* and through this stamp it becomes the *lex iuridica*.²⁰

Let us compare Kant’s thesis in the lecture with its corresponding thesis in the *Metaphysics of Morals*. The question in the lecture is part of legal method because it relates to an analysis of the application of the law. From this question the general legally theoretical viewpoint develops in the *Metaphysics of Morals* that a legal reality, with the legal character it has, is derived from the relevant law (the *lex iusti*) in the particular case.

¹⁹ Vigilantius writes “circumstantiarium,” but it should be “circumstantiarum.”

²⁰ On the details, see Byrd and Hruschka 2010, 52–58.

3 Reasons for excluding imputation

There are reasons to exclude imputation. In the lecture Kant lists some of these reasons (V 27:563.27–31). We can place these reasons in two groups. In the one group belong the circumstances under which a required act exceeds the power of the requirement's addressee or the requirement's addressee could not foresee the negative event, whose happening he could have avoided, or he was incapable of preventing the event from occurring. In the second group belong the cases in which the affected actor was "morally not permitted" or "had no authority" to avoid the undesirable event. In all these cases, as Kant says, the rule "*ultra posse nemo obligatur*" (no one has a duty beyond his capabilities) applies and the commission or omission of the act is "not imputable."

Let us consider an example for each of these groups. (1) A is drowning in a lake. C could save A if C were strong enough to pull A out of the water. C, however, is not strong enough. A drowns. A's death cannot be imputed to C. (2) A and B are drowning in a lake. C can save one, but not both. If he saves A then B will drown. If he saves B then A will drown. C, as stated in the *Metaphysics of Morals*, is in a collision of "reasons for an obligation." He has a moral reason to save A and he has a moral reason to save B but cannot act for both of these reasons. Let us, however, assume that B is C's minor son, whereby A is a complete stranger. In such cases the reasons for the obligation are not on equal footing, but instead one is "stronger" than the other, for which reason the weaker reason for obligation is subordinated (RL 6:224.9–26; cf. V 27:537.13–6). C is obliged to save B. Accordingly he is not permitted to save A. C saves B. A drowns. A's death cannot be imputed to C.

The "*posse*," or the ability or capacity to do something, in "*ultra posse nemo obligatur*," which Kant uses in the lecture to solve the cases he discusses, poses the same problem that we saw in discussing the concept of the *causa libera*. In the cases in the first group, the affected virtual addressee of the law (in the example, C) is not the author of the undesirable occurrence or nonoccurrence of the relevant event. The obligation he must fulfill, assuming the situation were different, he *cannot* fulfill. Accordingly, the *imputatio facti* is impossible. In contrast, in the cases in the second group, the virtual addressee of the law is indeed the author of the undesirable occurrence or nonoccurrence of the relevant event. That C in the example fails to save A is the result of his own *free* decision. The *imputatio facti* takes place. Instead, the law (in the example, the law that requires saving A) cannot be applied because C is "not permitted" to prevent the occurrence of the relevant event (the death of A) because C was obliged to

save B, which excluded saving A in the particular situation. Consequently, C was in this sense not “free” since he was prohibited from committing the relevant action (saving A). C *cannot* save A, but for different reasons than in the cases in the first group. It is not the *imputatio facti* that is excluded, but instead the *imputatio legis*. Kant’s use of the “*ultra posse nemo obligatur*” makes the “*posse*” ambiguous, just as the “*causa libera*” is ambiguous in the lecture.

The cases in the first group pose a further problem that Kant discusses in two places in the lecture. They concern the cases in which the “impotence to act or not act” came about because the actor was “*causa libera* of this impotence,” or in which the impotence “arose only through a condition” that lay in the actor himself (V 27:563.32–36 and 27:565.36–366.7). A classic case here is the drunk who commits an action (he hits and injures another person), which, at least in the initial analysis, cannot be imputed *because of his intoxication* since intoxication excludes the actor’s freedom. In the literature of the seventeenth and eighteenth centuries, two different solutions to this problem were proposed. (1) Some authors insist that the actor was not free at the time of the act and they impute only the causing of the impotence (in the example, the intoxication) if it was blameworthy. (2) Other authors accept as a *free* action (and thus as a deed) not only those actions that are free in themselves but also those actions that are not free in themselves but still, as they say, were free “*in their cause*.” Pufendorf is one of the latter authors. He only refuses to impute those actions “*quae penes ipsum neque in se, neque in sua causa fuit*” (which neither in themselves nor in their cause were within the power [of the actor]).²¹ Pufendorf’s Scottish commentator Carmichael, on the contrary, criticizes Pufendorf and adopts the first position indicated above (Carmichael 1739 [1718], 30 [Obs. 2 zu Lib. I Cap. I § 17]).

Pufendorf’s solution was adopted by a series of subsequent authors. At the beginning of the nineteenth century, the following terminology became common. The starting point is that all *actiones liberae in se* (all actions that are free in themselves) must be imputed. Imputable, however, are still such actions that are not free in themselves because they were committed in a defective condition, namely in those cases in which the actor was responsible for the defective condition. These actions are called *actiones liberae in (sua) causa* (actions free in their cause).²²

²¹ Pufendorf 1997 [1673], 16 (Lib. I Cap. I § 17).

²² On the history and for further references, see Hruschka 2003.

Kant changed his position over time. In the Kaehler lecture notes it says: “For imputation the action must evolve out of freedom; the actions of a drunk cannot be imputed but the intoxication itself can be imputed when the drunk was still sober (Kaehler 87.15–18). One can read something similar in Collins (C 27:288.23–26) and in Mrongovius (M I 27:1437.7–28). This approach is what we have listed earlier as “solution 1”. In the Vigilantius lecture, by contrast, Kant takes Pufendorf’s solution, solution 2. The passage at 27:563.32–36 is defective because the lecture refers to the imputability of the impotence (!) (and not the imputability of the action undertaken in a state of impotence). The passage at 27:565.36–566.7 is unambiguous. It says there: “Imputation occurs for every effect that is the result of a blameworthy caused ignorance or inability.” The action (here designated as “effect”) in the state of impotence can be imputed because it has its roots in ignorance or inability that the actor “blameworthy” caused. Kant brings the example that someone suffers from gout through excessive drinking. The gout makes him “unable to fulfill his official duties.” Here the person is imputed with his inability in office (and not only his drunkenness).

When Kant held his lecture in the winter semester 1793–4, *Religion within the Boundaries of Pure Reason* of 1793 had just appeared. There we find on this topic:

No reason in the world can stop him [i.e. the human being] from being a freely acting individual. One can correctly say, however, the consequences of his former free but unlawful actions can also be imputed to him; by saying this, one will say only: one need not accept this escape route and agree on whether the latter may be free or not because already in the admittedly free action that was the cause of these later actions there is sufficient reason for imputation. (RGV 6:41.10–17)

In *Religion*, Kant calls the action that was committed in a self-caused defective state a “consequence” of the causation of the defective state. The action is indeed itself an action but since it was committed in a defective state it is also a “consequence” of an earlier free action, namely of that action that caused the state. Imputation results if and because causing the defective state is “unlawful.” One who commits an action in such a state cannot rely on the action not being seen as free. We need not accept such an “escape route,” as Kant calls it. In other words: we need not consider whether the action committed in the defective state was free in itself or not because when the actor blameworthy caused the defective state, he committed a free action and that provides “sufficient reason” for the imputation of the action

committed in the defective state (the “consequence”). That is exactly the idea Kant has of the *actio libera in causa* in the *Vigilantius* lecture notes.

The Powalski lecture notes are interesting in this connection. There it says: “Nothing can be imputed other than what is free. If they are not *facta* then they are not *actiones liberae*. An action occurring against one’s will cannot be imputed but a free action, an *actio ab ignominia illicita*, can be imputed; this action is not directly a *factum* but indirectly” (P 27:156.24–28). This passage fits in our context if we assume the note taker omitted something from the Latin expression and assume that Kant in the lecture spoke of an “*actio libera ab ignominia illicita*.” An *actio libera ab ignominia illicita* is an action that is free because of a previous unpermitted ignominious deed. In other words: “*Actio libera ab ignominia illicita*” is another expression for “*actio libera in causa*.”²³ The final comment in the quote fits as well with this assumption, namely the action in question is not *directly* but *indirectly* a *factum*. In a first step the action in question is not free. Still if we view its negative history, then we must say that the action is a free action regardless of the defective state.

The Kaehler lecture notes come from the 1770s; Mrongovius I has the final date of February 1782; the Collins lecture notes are from the winter semester 1784–1785. The Powalski lecture notes, which seem to have been taken in 1777 or in one of the directly following years,²⁴ come from this period of time between the middle of the 1770s and the middle of the 1780s. It thus follows that Kant in the early period (in relation to 1793, the year of publication of *Religion* and the beginning of the *Vigilantius* lecture notes) wavered on the question, “Can we assume that an *actio libera in causa* was committed or not?” In Kaehler, Mrongovius I, and Collins, his answer is in the negative. In Powalski it is in the positive. Kant does not pick up the question again in the *Metaphysics of Morals*. From Kant’s point of view, he would have had to repeat the answer he gave in *Religion* and that would have been uninteresting.

4 Closing comments on imputation in the *Metaphysics of Morals*

Kant does remain interested in imputation and its problems in the *Metaphysics of Morals*. It is here one finds the often-quoted definition: “Imputation

²³ The “*illicitum*” in the expression “*actio libera ab ignominia illicita*” corresponds to “unlawful” (*gesetzwidrig*) in *Religion*, with which the basis of the defective state is more closely specified.

²⁴ On dating these notes, see, on the one hand, Lehmann 1979, 1043–1044 (on Powalski), 1050 (on Collins), 1052 (on Mrongovius I), and, on the other hand, Kuehn 2004, ix (in Stark 2004a).

(*imputatio*) in a moral sense is the judgment through which someone is seen as the author (*causa libera*) of an action, which is then called a deed (*factum*) and is subject to the laws" (MS 6:227.21–23). This is a definition of the *imputatio facti*. The reader will search for a definition of the *imputatio legis* in the *Metaphysics of Morals* to no avail. Still Kant does speak of "merit (*meritum*)" and "demerit (*demeritum*)" and thus of the consequences which he refers to as the "*imputatio legis*" in the lecture (e.g. MS 6:227.30–34).

An *imputatio legis* presumes that the applicable law is constitutive of a duty, and thus that one is speaking of a law of requirement or prohibition.²⁵ The concept of an *imputatio legis* thus is connected to a narrow concept of the practical syllogism, whose major premises are limited to requirements and prohibitions, such as in the example mentioned earlier (in Section 2) or in the following syllogism: "Lies are prohibited. – The action I am considering is a lie. – Therefore the action I am considering is prohibited." As a consequence, only those actions are susceptible of subsumption under this type of major premise that are required or prohibited. In the *Metaphysics of Morals*, Kant proceeds from a much broader concept of subsumption than in the lecture. The model he uses as relevant for the lecture is the *judicial* subsumption.²⁶ In the *Metaphysics of Morals*, not only the judge undertakes a subsumption. Instead, subsumption under a law is more of a daily practice of citizens who close contracts, for example, someone who buys rolls in a bakery, or more generally stated, someone who "by subsumption of a case" under the law "can acquire something or maintain what is his own" (RL 6:316.27–29). It is the citizen who applies the rules of private law, for example, those rules on acquiring property (the bakery rolls and other things). In comparison to the citizen's daily activities, a judge's activity is secondary.

²⁵ For this reason, as soon as permissive laws enter the picture there is no more sense in speaking of an "imputation of the law." Consequently, Kant says in the lecture "that all permissive laws lead to no imputation because the actions are *adiaphora*, and thus are not subject to duty or coercive law" (V 27:560.5–7). (The passage at V 27:562.5–7 contradicts this conclusion when it speaks of an *imputatio legis* also with a "*lex permissiva*," but this passage seems to be based on a mistake in the lecture notes.) Kant is obviously thinking here of permissive laws that make exceptions to requirements or prohibitions. Kant still discusses only permissive laws of this sort in *Perpetual Peace*; see ZcF 8:347.34–348.33, where Kant speaks of permissive laws being exceptions from laws of prohibition. We can consider the example of A injuring B in self-defense. One could apply the prohibition against battery here, but the situation of self-defense provides a justification and blame cannot be imputed to A. In a Reflection Kant comments on such cases: "If a person is found not guilty under the law then nothing can be imputed to him." Kant calls this nonimputation an "*imputatio negativa*," or a negative imputation (R #7155, 19:259.9f).

²⁶ The same is true of our modern understanding of the concept of subsumption, which is also oriented to the judicial subsumption.

Expansion of the concept of subsumption occasions a broader concept of imputation. Let us stay with our example of buying and selling bakery rolls. The seller and the buyer impute the activities they undertake to each other as legally relevant actions under the definition of imputation given above, and they do that *because* they see each other as persons.²⁷ They presume that the other contracting party is the author of his own actions, which are thus deeds (*facta*) and subject to laws, namely in our example the laws of buying and selling and of acquiring property. If they did not make these presumptions then the transfer and acquisition of property (rights to the rolls) would be inconceivable. We are *not* assuming that the externally exact actions (to those of buying a bakery roll) if committed by chimpanzees or robots would lead to the assumption (to *our* assumption) that the chimpanzees or robots have acquired or transferred property.²⁸ Kant distinguishes imputation which is done by a seller of bakery rolls as “evaluating imputation (*imputatio diiudicatoria*)” as opposed to a judge’s “legally binding imputation (*i. iudiciaria s. valida*)” (MS 6:227.23–26).²⁹ Against this backdrop, the cases of an *imputatio legis* form a (small) group of special cases that no longer need particular attention.

²⁷ Only persons are subjects whose actions are capable of imputation (MS 6:223.24f).

²⁸ On practical syllogisms in the *Doctrine of Right*, see Byrd and Hruschka 2010, 149–63.

²⁹ An “evaluating imputation” is not an abstract judgment but rather contains the contracting parties’ decision because that is the meaning of “diiudicare.” See entry “Diiudicare” in Heumann 1926 [1907], 147.

PART III

Ethics

Freedom, ends, and the derivation of duties in the Vigilantius notes

Paul Guyer

1 The essential end of mankind

In the lectures on moral philosophy that he gave before publishing the *Groundwork for the Metaphysics of Morals*, Kant is recorded as stating that “freedom according to a choice that is not necessitated to act” is “the inner worth of the world, the *summum bonum*” (C 27:344), that the “essential end of mankind” is that “as a free agent” a human being “must not be subject to his inclinations, but should determine them through freedom” (C 27:235), and that “the essential ends of mankind” (now in the plural, but for no obvious reason) are the “conditions under which alone the greatest use of freedom is possible, and under which it can be self-consistent” (C 27:346).¹ The essential end of mankind is thus that human beings set their own ends rather than having them imposed by inclination, and the fundamental principle of morality is that each use his or her own freedom to set ends subject to the sole constraint that such use be compatible with the freedom of all to set their own ends. In his 1784 course on natural right,² similarly, Kant is recorded as saying that “The inner worth of the human

¹ The close similarity of Collins’ notes to the earlier transcription by Johann Friedrich Kaehler from the summer semester of 1777, as well as the rather different opening lectures recorded by Carl Coelestin Mrongovius starting January 3, 1785 (see M II 29:597–639), strongly suggest that Collins copied his lectures or had them copied from an earlier source rather than taking them himself in class. Because the Collins transcription of Kant’s lectures on anthropology are dated 1772–1773 even though Collins was matriculated at Königsberg only a dozen years later, there is evidence that he purchased an older set of notes or had them copied for himself in another case. The Kaehler transcription of the lectures on moral philosophy is edited by Werner Stark (Stark 2004a).

² Gottfried Feyerabend inscribed his notes from this course “winter semester 1784,” but according to the catalog of the university in Königsberg Kant lectured on natural right in the summer semester of 1784. Either way, the course took place close in time to the lectures on moral philosophy through which Collins sat. Feyerabend’s notes on *Naturrecht* are included in the *Akademie* edition at volume 27:1317–1394; a new edition is being prepared by Heinrich P. Delfosse, Norbert Hinske and Gianluca Sadun Bordoni; thus far only the volume including the introduction has appeared: *Stellenindex und Konkordanz zum “Naturrecht Feyerabend,”* Teilband I: *Einleitung des “Naturrechts Feyerabend”* (Stuttgart-Bad Canstatt: Frommann-Holzboog, 2010). When that edition is complete, an English

being rests on his freedom, that he has his own will,” and that “If only rational beings can be ends in themselves, they can be thus not because they have reason, but because they have freedom. Reason is merely a means” (F 27:1319, 1321).

In the *Groundwork* itself, Kant does not repeat the formulation of the Collins lectures that freedom is the end in itself that is the inner worth of human beings or of the world, but instead uses a formulation like the second of those just quoted from the Feyerabend lectures in saying that it is “rational being” or its instantiation in human beings as “humanity” that is “an end in itself,” which lies “at the ground of a possible categorical imperative,” the most fundamental formulation of which must therefore be “So act that you use humanity, whether in your own person or that of any other, always at the same time as an end, never merely as a means” (G 4:428f). But several pages later Kant says that “Rational nature is distinguished from the rest of nature by this, that it sets itself an end” (G 4:437), thus implying that the rational nature or rational being or, in its instantiated form, humanity that is an, indeed the only, end in itself, is nothing other than the capacity of rational or human rational beings to set their own ends. Kant makes this implication explicit when, a dozen years later, he finally follows the *Groundwork for the Metaphysics of Morals* with the *Metaphysics of Morals* itself and defines humanity as that “by which alone” a human being “is capable of setting himself ends” or as “the capacity to set oneself an end – any end whatsoever” (MS 6:387, 392). This makes it clear that the humanity that is the end in itself is nothing other than the capacity to set ends, any ends whatsoever – subject to the condition, of course, as already made clear in the Collins lectures, that any use of this freedom be “self-consistent” (C 27:346) or that any use of this freedom by one person be consistent with the possibility of the further use of such freedom by that person on other occasions and by all other persons as well, insofar as their freedom to set ends might be affected by the other’s use of such freedom. And Kant makes the identity of the freedom that is the inner worth of mankind and through mankind the world with humanity even more explicit in the *Notes on the Lectures of Mr. Kant on the Metaphysics of Morals* taken by his friend and lawyer Johann Friedrich Vigilantius in the winter semester 1793–1794, thus three years prior to the appearance of the published work of that title, when he says that “Humanity is . . . thought of as pure intelligence in regard to the capacity for freedom and accountability

translation by Frederick Rauscher will appear in Kant, *Lectures and Drafts on Political Philosophy*, edited by Frederick Rauscher, translated by Frederick Rauscher and Kenneth Westphal, Cambridge University Press.

implanted in man” (V 27:579),³ again having defined freedom as consisting “only in this, that the agent utilizes his powers at his own choice,” although “in accordance with a principle of reason” (V 27:594). If humanity is the same thing as freedom and humanity is the end in itself, then freedom is the end in itself.⁴ Thus in finally executing his metaphysics of morals Kant comes back full circle to the original characterization of the fundamental principle of morality offered in his lectures on ethics two decades earlier, concluding that “An action is therefore *right or wrong*, only insofar as it accords or conflicts with the condition, that the agent’s freedom can coexist with that of anyone else, by universal laws, or is contrary to them . . . The reason for that is as follows: the universal law of reason can alone be the determining ground of action, but this is the law of universal freedom” (V 27:525).

With this background in place, the question to be addressed in this chapter can now be framed. In beginning his main discussion of duties of virtue in the *Vigilantius* lectures, Kant begins with perfect duties or *officia debita* to oneself, and states that the “duties of humanity” with regard to oneself “flow strictly, unconditionally and negatively from the concept of freedom” (V 27:601). That all duties, positive as well as negative, flow strictly and negatively from the concept of freedom is what we should expect him to say, given his statements that the fundamental principle of morality is simply the greatest possible use of freedom in accordance with reason’s requirement of self-consistency or universal law. The question to be addressed here is to what extent this statement is actually true of Kant’s derivation of all duties, including positive as well as negative duties to oneself as well as both duties of right and duties of virtue to others?. This is a problem because in the *Vigilantius* lectures Kant also says that in the case of duties of virtue, “Apart from the freedom of the action, there is . . . another principle present, which in itself is enlarging, in that, while freedom is restricted by the determination according to law, it is here, on the contrary, enlarged by the matter or end thereof” (V 27:543), and in light of things that Kant says in the published *Metaphysics of Morals*, it would be natural to read this as saying that duties of virtue depend on an end additional to that of freedom

³ Lehmann’s edition of the *Vigilantius* lectures in the *Akademie* edition, published in 1975, was based on an earlier publication of the original manuscript, now lost, by Emil Arnoldt (1907–1909). The *Vigilantius* lectures are structured around the same distinction between duties of right and duties of virtue or ethical duties that divides the subsequent *Metaphysics of Morals*, but it is striking that in the lectures Kant goes over the duties of virtue twice, from V 27:541–586 and 600–712. One can only wonder whether *Vigilantius* did not produce his manuscript from several antecedent sources.

⁴ Lara Denis has argued for a similar equation of Kant’s concepts of humanity and freedom in Denis 2010b, 171–173.

altogether, for example the happiness of others as an end entirely distinct from their freedom. Although such a line of thought is by no means alien to Kant, what I want to argue here is that the Vigilantius lectures in particular suggest an interpretation of this passage as saying that while negative duties to both others and oneself are duties to avoid the restriction of freedom, the principle of positive duties to others as well as to oneself is that of the expansion of freedom, and thus all classes of duty can in fact be derived from the concept of freedom combined with some basic facts about the human condition that bear on the realization of freedom.

2 Strategies for the derivation of duties

In the *Metaphysics of Morals* and the preceding Vigilantius lectures, Kant derives duties for human beings from the pure a priori and completely universal principle of morality valid for all rational beings that had been “sought out and secured” in the *Groundwork* and *Critique of Practical Reason*⁵ combined with the most fundamental facts about our embodiment and emplacement in the world. As Kant says, in the metaphysics of morals “we shall often have to take as our object the particular *nature* of human beings, which is cognized only by experience, in order to show in it what can be inferred from universal moral principles” (MS, Introduction, Section I, 6:217). In the works of the 1790s, Kant divides the specific duties of human beings into duties of right to others, which are a subset of our perfect duties to others, namely those that it is both physically and morally possible to enforce coercively,⁶ and ethical duties, which include our perfect duties to ourselves, our imperfect duties to ourselves and others, and those of our perfect duties to others that cannot be coercively enforced, such as the duties to avoid defamation and ridicule of others. Duties of right flow from the Universal Principle of Right, “Any action is *right* if it can coexist with everyone’s freedom of choice in accordance with a universal law, or if on its maxim the freedom of choice of each can coexist with everyone’s freedom

⁵ In Kant’s words, the “[g]egenwärtige Grundlegung ist aber nichts mehr, als die Aufsuchung und Festsetzung *des obersten Princips der Moralität*, welche allein ein in seiner Absicht und von aller anderen sittlichen Untersuchung abzusonderndes Geschäft ausmacht” (G 4:392). The derivation of the specific duties of human beings, in a way that adds to the fundamental principle of morality certain basic although empirical facts about the conditions of human existence, our embodiment and our emplacement on the finite surface of a sphere, falls under the category of the “other” or further “ethical investigation” to which Kant refers.

⁶ Kant does not make the two conditions on the enforceability of duties explicit, but the distinction I have invoked is standard in the account of perfect duties that he is appropriating from Gottfried Achenwall; see, for example, Achenwall and Pütter 1995, 61 (Book I, Chapter 4, §177).

in accordance with a universal law” (RL, Introduction, Section C, 6:230), that is, the principle that one may allow oneself as much freedom of choice and actions as is consistent with the equal degree of such freedom that one must concede to everyone else. It is easy to see how this principle can be derived from the “concept of freedom” or its universalization under the requirement of moral law, and how more specific duties for human beings can flow from it, for example the duty to establish a rightful system of property given that human beings need to use the resources of nature that could also be used by others and indeed need to use each other as resources for their own ends. It is also easy to see how perfect, negative duties to oneself, such as the duty not to destroy one’s own freedom by suicide nor compromise it by drunkenness or gluttony, could flow directly from the concept of freedom supplemented by the empirical but indubitable fact that human freedom of choice and action is dependent on the continued existence and condition of the living human body, and thus will be destroyed or impaired by the destruction or impairment of that body.⁷ It is less obvious how the *imperfect* duties to oneself and others, which Kant specifies as the duties to promote one’s own perfection and the happiness of others (and in at least one place identifies as the duties of virtue proper, as contrasted to the broader class of ethical duties that includes the noncoercively enforceable perfect duties to oneself)⁸ can be derived from the concept of freedom alone, or from the concept of humanity as an end in itself where humanity is equated with the freedom to set any ends whatsoever, even if that concept is supplemented by some basic facts about the human condition. Indeed, Kant frequently says that these duties are ends that are also duties, thus, as already suggested, that they depend upon the attribution of certain specific ends to human beings (see TL, introduction Section I, 6:380f), apparently ends *other* than the end of freedom itself, and that for that reason the derivations of these duties are synthetic rather than analytic (see TL, Introduction, Section II, 6:383). As Kant says in the *Vigilantius* lectures, “the *principia* of ethics are not to be derived from the nature of a man’s person,” that is, presumably, from the concept of humanity or the equivalent concept of freedom alone, “but must be unfolded synthetically, because the *officia meriti* [imperfect duties] can at all times be appended, merely, to the *officia debiti* [perfect duties]; in relation to the latter they are thus at all times ampliative, e.g., cultivation of talents,

⁷ Lara Denis has argued that “All duties in Kant’s system concern freedom and humanity” (Denis 2010b, 173), but has demonstrated in detail only how Kant’s perfect duties to oneself are derived from our obligation to preserve freedom in Collins, *Vigilantius*, and the Doctrine of Virtue of the *Metaphysics of Morals* (Denis 2010b, 174–181).

⁸ See TL, Introduction, Section II, 6:383.

promotion of the welfare of others" (V 27:600). In spite of this, my project here is to see how far we could get in deriving *all* the classes of duty from the concept of freedom alone plus the requisite factual assumptions about human nature that take us from the groundwork for the metaphysics of morals into the metaphysics of morals proper,⁹ and I will argue that the *Vigilantius* lectures themselves give us valuable guidance in this project.

There would seem to be at least four ways in which duties might be derived from the concept of humanity or freedom. First, and this is certainly the most natural interpretation of Kant's claims that duties of virtue proper are synthetic and ampliative rather than merely analytic, duties might arise from the combination of the moral law with the entirely *natural* ends of human beings, specifically from the *constraint* of the latter by the former. Second, and this might also be a natural interpretation of the claim that we have a duty to promote the happiness of others, thus to make their ends in a way our own, duties to promote ends might arise simply from the fact that those ends have been freely chosen on the assumption that the unconditional value of the free choice of human beings is *communicated* to the ends that they have freely chosen, at least insofar as those are consistent with the fundamental principle of morality, by a principle of transitivity of value. Third, duties might arise from the fact that humanity should not be equated solely with the capacity to set ends but rather with the capacity to set *and effectively pursue* ends: this would be a natural interpretation especially of the imperfect duty to cultivate our own talents. And, finally, all duties might somehow be nothing more than the forms of conduct necessary to preserve and maximize rather than minimize the freedom of human beings to set their own ends. I will argue that, although Kant indubitably employs the first method of deriving some of our duties, especially our imperfect duty to promote the happiness of others, that is not the only strategy he employs, especially in the *Vigilantius* lectures. But, I will argue, it is not the second or third strategies I have suggested by which he supplements the first strategy: Kant does not actually use them.¹⁰ It is the fourth, the idea that if not all of

⁹ In the *Groundwork*, Kant uses the term "metaphysics of morals" to designate the pure, completely a priori part of moral philosophy, its derivation of the fundamental principle of morality, which is valid for all rational beings and contains nothing empirical (see G 4:388); in the *Metaphysics of Morals*, however, he says that "a metaphysics of morals cannot dispense with principles of application" that take as their "object the particular *nature* of human beings, which is cognized only by experience, in order to *show* in it what can be inferred from universal moral principles" (MS 6:217). This is the sense of "metaphysics of morals" that is used in the title of the work, in parallel to the sense of "metaphysics" in the 1786 *Metaphysical Foundations of Natural Science*, and it is in that sense that I use the term here.

¹⁰ Although I and others have previously imputed the third strategy to Kant: Guyer 2005, 250, Guyer 2006, e.g., 254f; Johnson 2011, e.g., 94.

our duties then certainly more of them than might at first be imagined can be derived “strictly” from the concept of freedom as the forms of conduct necessary to preserve and expand freedom itself or humanity as nothing more than the freedom to set our own ends.

3 Deriving duties from freedom alone

Kant suggests one picture with a statement like this: “*en générale* . . . all duties of right, and the concepts to be formed of them, must be derived analytically from the concept of freedom, whereas all duties stemming from an end have to be demonstrated synthetically, merely, from the determination of human nature” (V 27:583). But he also says that “all obligation rests on freedom itself, and has its ground therein as far as freedom is regarded under the condition whereby it can be a universal law” (V 27:523). My question now is whether all of the particular classes of duty that Kant recognizes in his metaphysics of morals can be derived in accordance with the latter rather than the former statement.

The division of duties that Kant employs in the published *Metaphysics of Morals* is between duties of right, which are a subset of our perfect duties to others, and ethical duties broadly speaking, which include our noncoercively enforceable duties to ourselves, and our imperfect duties to both ourselves and others, the duties of self-perfection and of the promotion of the happiness of others, that is, the duty of love, which are also not coercively enforceable. Under imperfect duties to others Kant also includes what he calls the duties of respect, the duties to avoid arrogance, defamation, and ridicule, though it is not clear why, as negative duties or duties of omission, these are not perfect rather than imperfect duties – they call for the avoidance of certain behaviors without room for any exceptions or latitude – nor does Kant explain why these are not to be coercively enforced. In the *Vigilantius* lectures, Kant actually suggests a somewhat different classification, for instead of lumping our negative and perfect duties to ourselves together with our imperfect duties of self-perfection, though distinguishing the latter as duties of virtue within the broader class of ethical duties, Kant instead characterizes them as duties of right, though to ourselves and internal rather than to others and external, even though he also says that only the latter are “coercive, or genuine *officia juridica*,” while for the former “no external legislation is possible” but only an “inner one” or “self-coercion” (V 27:582, 587). This classification makes sense, because both the coercively enforceable perfect duties to others and the noncoercively perfect duties to oneself are duties not to destroy permanently or

temporarily the freedom of humans, whether oneself or others, insofar as that is compatible with the equal freedom of others, in accordance with the universal principle of right. “All rights are based on the concept of freedom, and are a result of preventing damage to freedom in accordance with law” (V 27:587), Kant says, and the perfect duties to self can be considered the “right of humanity in our own person” (V 27:592).¹¹

(i) It is easy to see how external duties of right or duties of right toward others are analytically derived or flow directly from the concept of freedom: the universal principle of right is, after all, the principle to claim only as much freedom for oneself as is compatible with universal law, which is to say always to allow others as much freedom as one claims for oneself. “An action is therefore *right or wrong*, only insofar as it accords or conflicts with the condition, that the agent’s freedom can coexist with that of anyone else, by universal laws, or is contrary to them” (V 27:525). But taken in its full generality, the principle of right also means that one must claim no *less* freedom for oneself than anyone could claim compatible with equal freedom for all others, in other words it is a principle to maximize not only the freedom of others but also of oneself, to the degree compatible with the maximization of freedom for all.

What does this mean when we interpret freedom or the humanity from which the right to freedom flows specifically as the freedom of each to set his or her own ends? In considering one case of wrongful action, one that could be enforced by law and thus be an *officia juridica* proper, Kant asks what would be wrong with not repaying a rich man “the money advanced, if is a burden to you, and he thereby incurs no significant loss?” He answers that acting on this maxim would be “a contradiction of the action with the law itself,” which would be a tautology if the law he meant were the law to repay debts; but obviously what he means is that action on such a maxim would contradict the principle of right, or the principle of allowing everyone, oneself included, the maximum freedom compatible with equal freedom for all. To see why this is so, one must presuppose the complement of what Kant adduces in the *Groundwork* as the principle of *hypothetical* imperatives: this is the principle that if one wills an end then, rationally, one must also will a sufficient means to it (G 4:417), and its complement is that it is rational to will an end only if some adequate means to it is available, otherwise one is not willing at all but merely wishing. Assuming this premise, then, refusing to repay the loan to the rich man would contradict the principle of right because if one’s proposed maxim were to be universalized, that is, freely adopted by all as it must be able to be if

¹¹ Denis also notes Kant’s discussion of perfect duties to self as a kind of duty of right in Denis 2010b, 178f.

it is to be rightfully adopted by you, then “nobody would lend to another” (V 27:532) and by willing that to be the case one would thereby have restricted one’s own freedom of action, that is, one’s own freedom to set ends, by depriving oneself of the ability rationally to set any end the only means to which would be a loan. Willing a maxim the universalization of which would contradict the possibility of acting upon it oneself deprives one of the ability to set certain ends for oneself, thus limits one’s freedom to set ends. The “right of humanity” in oneself as well as in others prohibits doing that.

But the more obvious cases of internal duties of right are those that damage one’s own freedom without obviously damaging the freedom of others. The first such duty that Kant considers, in both the lectures and the published text, is the duty to refrain from suicide and self-mutilation. These acts are prohibited because “[a] human being cannot dispose over his own substance, for he would then himself be the master over his very personality, his *inner freedom*, or humanity in his own person” (V 27:601). In other words, one cannot destroy the capacity to set ends that is located within one. But in the lectures more than in the published text, Kant also repeatedly emphasizes the prohibition of self-enslavement, because that would be a surrender of one’s ability to set one’s own ends. Thus he says that “[f]reedom consists only in this, that the agent utilizes his own choices, in accordance with a principle of reason; now anyone who ceded himself, with all his powers, to the disposition of another, and thus voluntarily enslaved himself, would alienate this freedom; he would treat his person as a thing, and this he cannot do” (V 27:594), and that one “cannot dispose over the causality of humanity, i.e., of freedom, insofar as this is *outer freedom*, in opposition to . . . inner freedom . . . He cannot therefore rob himself of his freedom, which would happen if he were willing to hand over the totality of his forces and powers for the arbitrary, absolute, unpermitted use of another” (V 27:601f). Kant’s contrast between “inner freedom” and “outer freedom” in the last passage might make it sound as if the person who would enslave himself to another might retain his ability to set ends and just be surrendering his ability to pursue effectively the ends that he sets for himself, and thus might seem to imply the prohibition of self-enslavement only on the basis of the bipartite definition of humanity as comprising both the abilities to set and to pursue ends that I previously said Kant does not actually employ. But because, again, one can rationally will an end only if one has adequate means to achieve those ends,¹² a person who gives up

¹² Or at least takes oneself to have those means; but a person in a condition of self-enslavement is hardly likely to be ignorant that he lost his powers to pursue his own ends.

control of his powers to another also destroys or restricts his freedom to set his own ends. That is what the right of humanity in one's own person demands one not do.

Other forms of conduct short of complete self-enslavement are also prohibited because they compromise one's freedom to set one's own ends. Of course, as we already know from the *Groundwork*, obtaining loans by false promises is prohibited, but Kant actually issues a more general injunction against borrowing on the ground that "To incur debts . . . involves a lowering in the value of one's own person, in that the debtor gives authority to the creditor to treat him arbitrarily (by extending the scope of his duty) . . . The debtor is put into the position, that with every call he will be expected to bring something" (V 27:605). In other words, by incurring debt one restricts one's freedom to set one's own ends and gives some of one's powers over to the creditor's freedom to set his own ends, whatever they might be. Kant likewise objects to begging because "[i]t is a human being's obligation to exert himself to the utmost to remain a free and independent being in relation to others; but as a beggar he depends upon the whims of others, and sacrifices his self-sufficiency" (V 27:605); in other words, the beggar fails to maintain to the maximum extent possible his freedom to set his own ends. And in yet another case that Kant frequently visits in the lectures, Kant characterizes miserliness as a violation of the right of humanity in oneself.¹³ "The miser merely collects resources, but eschews every method of employing them," or "possesses money, and thus a means of attaining all ends, but does not attempt thereby to become more perfect, since he renounces every use of it. He therefore contradicts himself in determining his own action" (V 27:606). This might make it sound as if in failing to use his money in any way that would increase his self-perfection, the miser violates his *imperfect* duty to self-perfection, which we are not yet discussing; but Kant's remark that the miser "contradicts himself in determining his own action" suggests that even more generally the miser is unnecessarily comprising his ability to set his own ends freely, thus compromising his own humanity. Kant suggests the same general point when he also writes that miserliness "consists in the desire to possess the means, without regard for the end that can be attained thereby" (V 27:659). Perhaps one could object that undertaking the project of collecting rather than expending resources is an end that one could freely set for oneself, but Kant's assumption is clearly that money is by definition merely a means to possible ends, and that by devoting oneself exclusively to

¹³ Kant gives more air time to the fault of miserliness in the lectures than he does in the published *Metaphysics of Morals*, where he devotes only one paragraph (§10) to the prohibition of avarice.

the accumulation of what is a mere means one is perforce restricting or surrendering one's capacity to set ends freely.

It should now be clear how the coercively enforceable duties of right toward others and even the noncoercively enforceable but still perfect duties of right toward humanity in oneself are duties to avoid destroying or damaging others' or one's own freedom to set ends. Let us now turn to the more controversial cases of the imperfect duties of self-perfection and happiness of others, which Kant himself has often suggested do not flow from the concept of freedom alone but has, as we saw, also suggested do so flow.

(ii) In the published *Metaphysics of Morals*, Kant explains why the perfection of others cannot be our duty without really explaining why our own perfection is. He then divides the objects of self-perfection between "cultivating one's *faculties* (or natural predispositions), the highest of which is understanding" and "the cultivation of one's *will* (moral cast of mind), so as to satisfy all the requirements of duty" (TL, Introduction, Section V.A, 6:387). In both the published work and the lectures, Kant lavishes attention on the cultivation of the moral capacities of self-knowledge and conscience, and his discussion of conscience in the lectures is particularly illuminating; but we can also say that our duty to cultivate these faculties is second-order, that is, we need to cultivate them in order to be able properly to fulfill all our first-order duties, or at least our first-order ethical duties, where no external coercion is available, but they do not by themselves tell us anything about the contents or source of those first-order duties. So we need to see in particular how the duty to cultivate our natural rather than moral predispositions derives from our duty with regard to our own humanity or capacity to set ends directly, if it does. We might also note that in the published work Kant first includes the duties to cultivate self-knowledge and conscience among the perfect duties to oneself (§§ 13–15), which are otherwise the negative duties not to destroy, injure, or stupefy oneself or to lie or fall victim to avarice, but then again includes the duty to "increase" one's moral perfection among the imperfect duties, not on the ground that this duty is not entirely determinate but on the ground that human beings can only strive for moral perfection but not attain it, at least in this life (§§ 21f).¹⁴ This repetition is confusing, since it

¹⁴ Thus Denis discusses the duty to *moral* self-perfection, since Kant includes that under perfect duty to oneself, but not the duty to the perfection of one's *natural* capacities, as Kant never describes that as a perfect duty (Denis 2010b, 176). It might also be noted that the duty to cultivate the conditions necessary for the fulfillment of duty actually blur the line between duty to self and duty to others: for example, the duty to cultivate one's natural disposition to sympathetic feelings that Kant describes in TL, §§ 34f, is a duty to cultivate something *in oneself for the benefit of others*, so might well seem to be a duty to self and to others at one and the same time.

is not clear that one's moral perfection consists in anything other than the perfection of one's self-knowledge and conscience, or even if the perfection of the moral will includes something over and above these it is not clear that it is not imperfect in precisely the way described. In the lectures, this confusion is not present. As we have already seen, in the lectures the negative duties entailed in not destroying or damaging one's own freedom and thus one's own humanity have also been hived off as internal and noncoercively enforceable duties of right, so they are not dealt with under the rubric of self-perfection. All of this leaves the duty to perfect our various natural capacities, divided in the published *Metaphysics of Morals* into the duties to cultivate our "spirit," or understanding and reason, our "soul," or supporting intellectual powers such as memory and imagination, and our bodily strength and facility (§ 19), to be derived from the concept of freedom itself.

There are passages in the lectures in which Kant derives the duty for the perfection of one's not directly moral capacities *instrumentally*, as the perfection of the means needed for the promotion of the happiness of others and oneself. This passage suggests that thought:

In regard to ourselves, the maxim of duty can be directed only towards rendering perfect, since to further the happiness of others is an end, the means to which I can furnish no otherwise than through my own perfection, in order to act in accordance with this moral aim. For perfection as such, and taken as *genus in abstracto*, is completeness, suitability of a thing to all kinds of ends, or formal perfection in relation to every material perfection that one can enumerate singly in regard to all the capacities of mind or body. *In specie*, here, in the moral sense, perfection is the conformity of all our powers with the end of humanity, i.e., happiness; and if our actions are directed to seeking our own perfection for the happiness of others, they conform to the end of humanity. (V 27:651)

However, another passage suggests a very different argument. Here Kant says that:

In regard to our own person, humanity is an ideal, to which we owe the duty of perfecting ourselves, so that we may fulfill the duties that it imposes on us. Now cultivation itself is applied to an inner capacity of the soul, by whose essential nature man attributes a free person to himself, and which is thus the personality of man, as a being endowed with freedom. So all duties that are incumbent on him in regard to his own person take account of him as noumenon, or as a being that acts in freedom. (V 27:626)

There is no mention of happiness here, only of freedom, so the duties that are to be derived from "humanity as an ideal" must be duties concerned with the preservation or promotion of freedom.

Kant then continues that the duties that follow from this basis “can be reduced to three general determinations,” namely, first, “*To possess oneself*, i.e., to determine all actions by way of a free choice . . . For man has a capacity to employ himself in a purposive way [*zweckmäßig*]. But he attains this only by subjecting all his powers and capacities solely to his free choice, and employing them accordingly” (V 27:626); second, “*The duty to govern oneself*,” which “involves cultivation of the mental powers to those ends with which they are collectively compatible, and constitutes, therefore, the essential in the soul’s capacity or readiness to enlarge the *facultates animi* for all moral ends” – and then there is a gap in Vigilantius’s notes, and the third “determination” is not explicitly mentioned. And when the text resumes, Kant is discussing the impermissibility of suicide, which is of course a duty of self-preservation rather than self-perfection, so there is obviously a missing transition and there is no way to gauge how much material is missing. But in his discussion of the two “determinations” that our notes do include, Kant states that the duty of perfecting our (nonmoral) capacities of spirit, soul, and body is necessary in order to *determine all our actions by way of free choice* or to employ all our *capacities in a purposive way*; in the second determination, he says that we must enlarge the faculties of our soul for *all* moral ends, but these phrases in the first determination suggest that we must enlarge the faculties of our soul for all ends, without restriction. And what can this mean? Not that we must cultivate our faculties in order to be able to achieve our ends, as he says later, but that we must cultivate them in order to *expand rather than contract the scope for our choice of ends*. Why would this in turn be so? Because, again, by the general principle of reason according to which we can rationally will only those ends for which we have adequate means, failing to cultivate our natural predispositions to talents and capacities limits our possible choice of ends, while cultivating those predispositions, precisely by expanding the means available to us, also expands the range of ends we may rationally choose. In other words, perfecting what might seem like mere means to ends already chosen or given in fact enhances our freedom to choose ends, or perfects our freedom itself. The freedom to set ends might seem like something entirely internal, as the opening of the passage presently under discussion seems to say, entirely independent of the question of whether means to realize those ends are available in the external world, but freedom is not independent of external conditions in that way; rather the scope of our freedom of choice depends on the availability of means to the ends we might want to be able to choose.

The principle of the imperfect positive duty of self-perfection is thus, on this approach, the complement to the perfect negative duty not to destroy

or limit one's own freedom of choice: just as miserliness, for example, sacrifices the freedom to set ends to its obsession with the accumulation of means, self-perfection develops our natural capacities as means to expanding our freedom to choose our own ends. On that note, we can now turn to the question of whether our imperfect positive duties toward others can also be derived from the concept of freedom rather than from that of happiness.

(iii) Of course Kant explicitly describes our positive duty of virtue or "duty of love" toward others as the duty to promote their happiness in the *Metaphysics of Morals* and as we have already seen frequently describes it in the same way in the preceding lectures. This is a duty that arises from our own natural desire for or end of happiness, the natural desire of others for their own happiness, and the requirement that we act only on universalizable maxims and thus seek or even hope for assistance from others in the pursuit of our own happiness only if we are willing to assist them when an appropriate occasion arises. The freedom of the others whose happiness we are to promote does not seem to play a role in this argument. There is, however, a role for freedom built into even this conception in the form of Kant's requirement, stressed in the Introduction to the Doctrine of Virtue and also expressed in his antipaternalist theory of the state, that we promote the happiness of others only in accordance with *their own* conceptions of happiness, not in accordance with *ours*: "It is for them to decide what they count as belonging to their happiness," although it is also "open to me to refuse them many things that *they* think will make them happy but I do not, as long as they have no right to demand them from me as what is theirs" and, presumably, I also have other ways in which to fulfill my imperfect duty to promote the happiness of others – which, given the nature of the human condition, I surely always will.¹⁵ But this says only that preservation of the other's freedom to choose his own ends is a necessary condition on the promotion of happiness, as of course it must be if the preservation of freedom is the principle of our perfect duties to others as well as to ourselves and perfect duties always trump imperfect duties. It does not say that the duty to promote the happiness of others itself derives from an imperfect duty to promote or expand the freedom of others.

And it would be difficult to find an explicit statement of the latter in the *Vigilantius* lectures, because in fact these lectures say virtually nothing about the general duty to promote the happiness of others beyond what

¹⁵ See TL, Introduction, Section V.B, 6:388; for Kant's political antipaternalism, see, for example, TP 8:290f.

we have already seen. Instead, where we might expect a discussion of this general duty, the lectures turn almost immediately to our imperfect duties to others who stand in particular relations to us, especially our parents, our children, and our friends, his discussion of which Kant in turn connects to his discussion of our duties with regard to our spouses. This is striking, because while in the *Metaphysics of Morals* Kant discusses our duties to spouses and children under the rubric of acquired private *right*, his placement of his discussion of these relationships under the category of our imperfect duty to others in the Vigilantius lectures suggests that while of course we have a duty to promote the happiness of mankind in general, in fact the characteristic objects for the fulfillment of our imperfect duty to others are those to whom we stand in special, indeed intimate relationships: it is in our relationships to those close to us that most of our moral lives occur, not in our relations to strangers. And in Kant's discussion of these relationships in the Vigilantius lectures, promoting the moral end of freedom is indeed more clearly the source of our duties than promoting the merely natural end of happiness.

Thus, Kant writes about parental duty toward children that "the rearing of children up to the point of self-sufficiency, i.e., an education so ordered that the children thereby obtain contentment with their lot, and pleasure in their existence, is an *opus supererogationis* on the parent's part, a thing superadded, a kindness which involves something meritorious" (V 27:670). This is a remarkable statement in several ways. First, Kant's claim that the duty of education is supererogatory might seem implausible: although it arises from his premise that perfect duties are always negative duties not to harm while imperfect duties, the fulfillment of which is meritorious, are positive duties to benefit, the line between not harming children by withholding some reasonable amount of education and positively benefiting them by providing an adequate or more than adequate education seems fuzzy, to say the least. But what seems most important in the passage is its suggestion that the aim of education is to make children capable of self-sufficiency, and that their happiness is to result from their self-sufficiency. Throughout Kant's discussion of duty to self, self-sufficiency is closely linked to and one of the forms of freedom: self-sufficiency consists in not allowing oneself to become dependent on external objects and on others, thus retaining the ability to set one's own ends (V 27:652f), but also in developing the means to satisfy possible ends and thereby enhancing one's freedom to set one's own ends. This is what parents (and no doubt other adults, ranging from grandparents and guardians to teachers to taxpayers) must help children develop, and then, Kant suggests, the happiness and

contentment of the children will follow, because contentment lies in exercising one's own freedom and happiness follows not just from having one's ends satisfied but in satisfying ends one has chosen for oneself. In other words, our fundamental imperfect duty to children is to enhance their freedom, and the enhancement of their happiness is primarily a consequence of that.

In the lectures, Kant immediately follows his discussion of our duty to educate our children with a discussion of friendship. Kant uses the word "freedom" only once in the discussion of friendship that concludes the "elements of ethics" in the published *Doctrine of Virtue* (§§ 46–47), and does not use it at all in his extended discussion of the topic in the *Vigilantius* lectures (V 27:675–686). In the latter discussion, he begins by characterizing friendship as "reciprocal" "well-wishing love to others" (V 27:675f) and continues by characterizing the "reciprocal enjoyment of humanity" that takes place in friendship as "mutual relation in regard to capacity, and the satisfaction of the power and need so typical of human being[s] to stand together, to communicate not only their feelings and sensations to one another, but also their thoughts" (V 27:677). Here Kant does not explicitly connect the concept of humanity to the capacity to set ends, though no doubt it is implicit in the concept of mutual well-wishing that friends want to promote the freely chosen ends of each other because those are freely chosen ends. This is also implicit in Kant's requirement that "Those who love" in friendship be "equals; well-wishing, love and friendship differ from *favour* in this respect. The ability, that is, to promote the other's well-being, and do him good, must be the same in both . . . *inter superiores et inferiores* no friendship occurs" (V 27:676). The reason for this restriction is that when a superior confers a benefit on an inferior, the freedom of the latter is restricted in a way that the freedom of the former is not: the inferior will feel the need to return a favor he cannot afford, or that will force him to give up some other end in a way that the original conferral of benefit by the superior did not force *him* to give up anything equivalent, and so on; in some way or other, the inferior will suffer a loss of freedom that the superior did not. But perhaps most important, Kant's rationale for the importance of sharing thoughts and feelings in friendship is ultimately that this enhances the freedom to set ends of each friend. Kant writes:

For the possession of a friend affords pleasure only in that the need to communicate our feelings and thoughts can be satisfied quite unconditionally and unrestrictedly, without reserve or any seeking for advantage, and thus without interest. Only that pure interest which is the end of this inclination converted to a need, namely to perfect our acquirements

[*Kenntnisse*] by communicating, rectifying and determining them through the judgment of others, is the one pure goal that must lead us to friendship. This is all the more beneficial, in that we cannot rectify circumscribed ideas and thoughts in any other way than by sharing them, and should this not occur, we are never secured against errors. (V 27:683)

Our “pure interest,” at least according to Kant’s moral theory as a whole, can only be to preserve and promote our own freedom and that of others to set ends. What this passage then argues is that to do this as best we can – and thus to maximize our freedom to set ends – we need our judgments – judgments about what ends would be best for us, what ends would be most moral, that is, most compatible with and beneficial for the ends of others as well, and what means would be most effective for our ends – to be as well-informed as they can be, and that reaching such a level of clarity is not something we can do on our own, but something we can do only through frank and open discussion with people we respect and trust to have both our own interests and those of mankind at large at heart. That is what friends are for: to enhance the freedom of each but also to assist each in enhancing the freedom of all.

Kant also suggests such a conclusion in his striking comparison of friendship to marriage. Kant’s view of marriage is much too large a subject to address adequately at this late stage of the present chapter,¹⁶ but one point will suffice for present purposes. In his discussion of friendship in the *Vigilantius* lectures, Kant says that “It is somewhat the same” in friendship “as in marriage,” namely each is “communal possession of one person by the other, or reciprocal possession, i.e., union of their person as to moral disposition . . . each mutually shares in every situation of the other, as if it were encountered by himself; and this, indeed, by laws of moral freedom” (V 27:677). This is very general, but what Kant stresses in his discussion of marriage in the *Metaphysics of Morals*, where it comes in the Doctrine of Right, not the Doctrine of Virtue, is that marriage is the only way to avoid treating oneself as well as another as a mere means to pleasure in sex, a mere thing, because although in marriage it might initially seem as if “one person is acquired by the other *as if it were a thing*,” nevertheless, since “the one who is acquired acquires the other in turn . . . , in this way each reclaims itself and restores its personality” (RL § 25, 6:278). Because the contrast to using oneself and the other as a mere means is treating both oneself and the other as an end, and that in turn means allowing each to set his or her own ends as freely as possible consistent with the equal freedom of the other, this must mean that

¹⁶ For valuable discussions, see Herman 1993 and Ebels-Duggan 2008.

marriage does (or can) turn sex into an exercise of freedom: by extending freedom to the other to set his or her own ends in having sex, the other controls his or her own sexual inclinations by this moral concession, and thereby transforms his or her own sexual conduct from a mere gratification of inclination into an exercise of freedom. Marriage is thus a vehicle for transforming acting on inclination into exercising freedom. It is this feature of marriage that can also be found in friendship, where for example the selection of pastimes or a place for dinner that might otherwise be a matter of mere inclination, although presumably less fraught than the choice of sexual practices, can still be transformed into an exercise of freedom of each by the concession of freedom to the other that underlies any agreement achieved through mutual communication and deliberation.

This discussion of friendship and marriage allows me to conclude with a point about Kant's distinction between duties to self and others, and to some extent about his distinction between duties of right and of virtue as well: namely, that these distinctions are abstractions, and that in real life how we treat ourselves and how we treat others, as well as how we preserve freedom and how we enhance freedom, are typically intertwined. In a friendship and in a marriage, we typically enhance our own freedom and that of the other at the same time, thus fulfilling duty both to self and other at the same time. And in marriage but presumably in friendship as well, we both preserve or protect and promote and enhance the freedom of both ourselves and others at the same time. Furthermore, the cultivation of our natural predispositions to skill, knowledge, and so on is necessary both to protect and to enhance the freedom of both ourselves and others; it not merely a means to achieving ends that are independently and antecedently set, but a necessary condition for preserving and enhancing our freedom to set ends. As Kant shows in some detail in the *Vigilantius* lectures, we do not need to bisect the definition of humanity into the separate capacities to set ends and to pursue them effectively to derive the whole duty of man,¹⁷ but can derive it from an adequate understanding of what the freedom to set ends really involves.

¹⁷ Or two of the three parts of what Pufendorf called the whole duty of man, that is, duties to self and to others; for Kant, of course, we have no duties to God.

Proper self-esteem and duties to oneself

Lara Denis

From early lectures to late published works, Kant links self-esteem [*Selbstschätzung*] and duties to oneself. The Collins notes on Kant's lectures on moral philosophy identify self-esteem as the "*principium* of the self-regarding duties" (C 27:347). The *Tugendlehre* describes self-esteem as "the basis . . . of certain actions that are consistent with [one's] duty to [oneself]" (MS 6:403). The Collins and Vigilantius notes as well as the *Tugendlehre* portray violations of duties to oneself as fundamentally opposing an agent's self-esteem, rendering her contemptible – worthless – in her own eyes (see C 27:341f, 343; V 27:605; MS 6:430). Kant's association of self-esteem with duties to oneself merits attention for what it might reveal about duties to oneself, to which Kant attributes primacy (see C 27:341, 433; V 27:579f, 604). It requires investigation, for "*Selbstschätzung*" refers to more than one thing.

This chapter aims to elucidate Kant's conception of self-esteem – particularly in its morally correct forms – and its relation to duties to oneself. I concentrate on the Collins and Vigilantius lecture notes and the *Tugendlehre*. The last contains Kant's mature, published view of these matters. The lecture notes include lengthier discussions of self-esteem, present different accounts of duties to oneself, more clearly show where Kant is responding to Baumgarten, and provide a window into the development of Kant's thought. The notes also contain Kant's more colorful claims about the baseness of violating duties to oneself (see C 27:343f, 373; V 27:604).

Section 1 delineates several items Kant discusses under the heading of "*Selbstschätzung*," explaining some of their relations to duties to oneself and one another. **Section 2** explicates duties to oneself as acts that instantiate the moral attitude of self-esteem. **Section 3** considers the relation of self-esteem to the primacy of duties to oneself; here I interpret Kant's dramatic claims that by violating duties to oneself one throws away one's humanity, loses all inner worth, and makes oneself contemptible, as attempts forcefully to convey the primacy of duties to oneself by appealing to his audience's

feelings and desires concerning their own worth – essentially, that is, to their self-esteem.

1 Selbstschätzung

Kant identifies *Selbstschätzung* in its morally correct forms with a way of estimating one's worth; an opinion (or set of opinions) of one's worth; a feeling of worth; and a principle and an attitude concerning one's worth. Examination of each individually points toward a general conception of proper self-esteem as a morally correct way of valuing oneself in which all of these items are elements.

1.1 Activity

One referent of “*Selbstschätzung*” is the activity or practice of estimating one's inner worth by the standard of the moral law. In *Vigilantius*, where he gives the most systematic account of it, Kant says, “a just self-estimation [*rechtmäßige Selbstschätzung*], the *justum sui aestimium*,” “consists in a judgement of our moral worth, i.e., in a testing of our action by its agreement with, or deviation from, what is said by the moral law, and by the extent to which it is undertaken, not merely in accordance with, but for the sake of the law alone” (V 27:609, translation modified).

Although we human beings are naturally inclined to compare ourselves with one another, and although interpersonal comparisons are appropriate in some contexts, other human beings provide an incorrect standard for assessing our moral worth (see V 27:703). The moral law does not determine our duties relative to what human conduct indicates about our capacities (see MS 6:404f). The standard presented by others is “contingent” (C 27:359): we may fare well or poorly depending on whom we compare ourselves with.¹ Even the best people are “but copies” of the law (C 27:349). The standard of other human beings is irrelevant: whether others are better or worse than we are has no bearing on our worth (see H 27:40). Finally, interpersonal moral comparisons encourage such vices as envy and *Schadenfreude* (see C 27:359, 436–440).²

In *Vigilantius*, Kant depicts self-estimation as having two aspects. We must ascertain both “the worth of humanity in our own person, i.e., the worth which can be laid upon the actions by the intelligible human being,”³

¹ Compare C 27:436 with Smith 1969 [1759], 362f.

² See Smit and Timmons in this volume.

³ Throughout, I modify Heath's translation by rendering “*Mensch*” as “human being.”

according to the whole determination of his existence,” and “the worth of the human being in [our] person” (V 27:609). We properly estimate the latter by comparing it with the former, for the moral law “in respect of its definition is identical with humanity and the Idea thereof” (V 27:610; see also MS 6:480).

Just self-estimation of the human being in our person must be strict and impartial. We must avoid both exaggerating our conformity to the moral law and minimizing the moral law’s stringency. The second error is worse, for it is a corruption of the standard rather than its application (see C 27:348). To say we must be strict does not mean we may inflate our faults: “we must not render our condition either better or worse by fabrication; in either case, the judgement about moral worth as such then comes out wrong” (V 27:608).

Finally, self-estimation is not supposed to be a one-time act, or focused exclusively on the present. It is an essential component of the ongoing quest for moral self-knowledge and perfection (see V 27:608f; C 27:348; MS 6:441f).⁴ When we engage in it, we must consider not only the morality of our current conduct, but our successes and failures adhering to past resolutions for improvement (see V 27:608; RGV 6:68f). Striving for self-knowledge, to which just self-estimation is integral, is presented in Vigilantius and the *Tugendlehre* as the “first” duty to oneself (see V 27:608⁵; MS 6:441f). It “counteracts that *egotistical* self-esteem, which takes mere wishes ... for proof of a good heart” (MS 6:441). Yet it also illuminates the dignity of humanity in our person and the sublimity of our moral vocation, thereby averting despondency and “fanatical” self-contempt.

1.2 Opinion

As an *opinion* of ourselves regarding our worth in comparison with the moral law, *Selbstschätzung* may be high or low, depending on its object: “We have reason to harbor a low opinion [*Meinung*] of our person, but in regard to our humanity we should have a high opinion” (C 27:348f, translation modified). Kant identifies the low opinion of oneself as human being with “humility,” the high opinion of humanity in our person with “noble pride” (C 27:348).

⁴ In *Ethica* §168, Baumgarten contrasts humility with “*iustum sui aestimium*,” which in the 1763 edition he translates as “*gehörige Selbsthatzung*” [*sic*]. He refers to each as a “*habitus*”: humility as the habit of correctly assessing our imperfections, and proper self-esteem as the habit of correctly assessing our perfections (AA 27:779, 914).

⁵ The Vigilantius notes make explicit that Kant is following Baumgarten here.

Humanity in our person is congruent with the moral standard; the human being is not: “Humanity itself, if we wish to personify it, actually lacks any inclination to evil, but the more a human being compares himself therewith, the more he finds out how far away he is from it” (V 27:609). “The human being has reason to have a low opinion of himself, since his actions are not only in contravention of the moral law, but also lacking in purity” (C 27:350; see also KpV 5:78). Various individuals have reason to think better or worse of themselves morally, based on their conduct and what it suggests about their disposition (KpV 5:37f, 88, 123; RGV 6:68f). Still, as “in human beings all good is defective” (KpV 5:77), our own hearts are “inscrutable” (RGV 6:51; see also G 4:407), and we are prone to flattering self-deception concerning our worth (see RGV 6:68), even the most virtuous person has reason to have a low opinion of herself in comparison with the law.

No one, however, should have a low opinion of her moral worth as a human being in comparison with others. We have the same capacity for virtue, through which we can gain inner worth: “in comparison with others, we have no reason to entertain a poor opinion of ourselves, for I can just as well possess worth as anyone else” (C 27:349, see also 359; V 27:609f). Presumably, some people are morally better than others (see C 27:351, 462). Yet we cannot – and should not attempt to – estimate the inner worth of others, or compare it with our own (see C 27:450f; V 27:695, 703f). Furthermore, the humanity in each of us is the same, so we should regard ourselves as equal to one another in regard to the dignity and worth of humanity in our persons.⁶

1.3 *Feeling*

Sometimes Kant identifies *Selbstschätzung* with a feeling of high inner worth based on our autonomy and moral capacity. Among the three works on which we are focused, the *Tugendlehre* alone contains a developed account of this feeling. Kant says that humility, as “a consciousness and feeling of the insignificance of one’s moral worth,” follows from comparison with the moral law (MS 6:436; see also KpV 5:154).

But from our capacity for internal lawgiving and from the natural human being’s feeling himself compelled to revere the (moral) human being within

⁶ The merit one human being may have in relation to another entails neither a disparity in personal worth between them, nor a difference in their fundamental moral standing. On merit in relation to others, see V 27:558, 611, 665; MS 6:455.

his person, there comes *exaltation* of the highest self-esteem, the feeling [*Gefühl*] of his inner worth (*valor*), in terms of which he is above any price (*pretium*) and possesses an inalienable dignity (*dignitas interna*), which instills in him respect for himself (*reverentia*). (MS 6:436)

Kant equates this feeling of self-esteem with “the human being’s feeling for his sublime vocation” and an “*elation of spirit (elatio animi)*” (MS 6:437).

As a susceptibility to pleasure in the awareness of our freedom, this feeling is present in all human beings.⁷ In the Introduction to the *Tugendlehre*, Kant identifies self-esteem with “*respect for oneself*” and as one of four “aesthetic preconditions of the mind’s receptivity to concepts of duty as such,”⁸ which “lie at the basis of morality as *subjective* conditions of receptiveness to the concept of duty” (MS 6:399). Our consciousness of them follows “from consciousness of a moral law, as the effect this has on the mind” (MS 6:399). Kant here describes self-esteem as a feeling “of a special kind” which “is the basis of certain duties, that is, of certain actions that are consistent with [a human being’s] duty to himself” (MS 6:403). There is no duty to have this feeling; rather, “the law within [the human being] unavoidably forces from him *respect* for his own being” (MS 6:402f): “It cannot be said that he *has* a duty of respect for himself, for he must have respect for the law within himself in order even to think of any duty whatsoever” (MS 6:403).

The second *Critique* clarifies the connection between respect for the law and self-esteem. In the “Incentives” chapter, Kant presents the moral law as an incentive (respect) as affecting the mind in two ways. We experience *humiliation* as the moral law strikes down self-conceit [*Eigendünkel*] – an unwarranted claim to esteem for ourselves – removing “counterweights” to the law (KpV 5:73f). Yet with this humiliation comes *elevation*, for this supremely authoritative law springs from our own pure practical reason (KpV 5:79–81; see also RGV 6:23n.25–30). In our determination solely by the moral law, we are conscious of our freedom and feel a “self-approbation in reference to pure practical reason” (KpV 5:81; see also G 4:449). Thus, our esteem for ourselves as free is inextricable from respect for the moral law.

The first sentence of the concluding paragraph of the “Incentives” chapter emphasizes that the feeling of self-esteem is integral to human moral motivation:

This is how the genuine moral incentive of pure practical reason is constituted; it is nothing other than the pure moral law itself insofar as it lets us

⁷ On feeling, see MS 6:211f.

⁸ Translation modified following Guyer 2010, 137f.

discover the sublimity of our own supersensible existence and subjectively effects respect for their higher vocation in human beings, who are at the same time conscious of their sensible existence and of the dependence, connected with it, on their pathologically affected nature. (KpV 5:88)⁹

The *Tugendlehre* account builds on this earlier one by portraying morally motivated compliance with self-regarding duties as depending on the feeling of self-esteem in a uniquely immediate or fundamental way. Calling the feeling of self-esteem a “basis” [*Grund*] for actions that agree with one’s duty to oneself (MS 6:403)¹⁰ suggests that moral motivation to perform self-regarding duties operates directly through this feeling (see Guyer 2010, 150f).

The second *Critique*’s “Doctrine of Method” associates our feeling of worth in following the moral law with our dread of finding ourselves “worthless and contemptible” in our own eyes; this dread facilitates our development of virtues and hinders our development of vices (KpV 5:161). The *Tugendlehre* account suggests that this dread is especially conducive to the development of self-regarding virtues and especially antagonistic to self-regarding vices. Both works suggest that, though there is no duty to have the feeling of self-esteem, there is a duty to strengthen it.¹¹

1.4 Principle

In Collins, Kant calls *Selbstschätzung* the “*principium*” of duties to oneself: “The *principium* of the self-regarding duties does not consist in self-favour, but in self-esteem; our actions, that is, must be in keeping with the dignity of humanity” (C 27:347, translation modified).

We might reasonably take this *principium* to be an objective moral principle which grounds the obligation to duties to oneself as acts necessary for compliance with it.¹² So understood, it seems roughly to anticipate the self-regarding component of the formula of humanity: “*So act that you use humanity, whether in your own person or in the person of any other, always at*

⁹ A fragment from 1776–1778 anticipates Kant’s mature view about self-esteem and moral motivation. See R #6866, 19:185f.

¹⁰ Kant’s language might indicate that Kant takes this feeling to underlie a wider or narrower class of actions than all and only those enumerated as duties to oneself in the *Tugendlehre*’s “Elements of Ethics”: for example, all ethical duties (MS 6:418; V 27:579) or only those actions required by the “duty of self-esteem” (MS 6:462).

¹¹ On its cultivation, see MS 6:483; KpV 5:152f; on dangers to it, see KpV 5:156f; KU 5:273.

¹² Kant presents the Ulpian principle *honeste vive* as a fundamental principle of self-esteem – and as the basis of all duties to oneself and perhaps of all ethical duties (C 27:280f; MII 29:631f; V 27:527, 668; MS 6:236). Despite its relevance and interest, I cannot examine this principle here. See Byrd and Hruschka 2010, 62–67; Höffe 2010, 85–87; and Ripstein 2009, 18.

the same time as an end, never simply as a means" (G 4:429; see also 436, 438).¹³ Humanity here is our freedom; we are ends in themselves in virtue of our freedom (see G 4:448f; F 27:1322; KpV 5:60f). In Collins, Kant presents *acting in keeping with the dignity of humanity* as equivalent to *conforming behavior to the essential ends of humanity (or human nature)* and *restricting our use of freedom to agreement with the conditions of its fullest self-consistent expression* (C 27:345f). Thus, the principle Kant identifies with self-esteem is a principle of constraining our freedom only by itself, not subordinating it to inclination. Kant identifies self-esteem not only as "the general *principium* of duties to oneself" but also as "the objective condition of morality" (Kaehler 202; cf. C 27:360);¹⁴ one acts wrongly if one's use of freedom is inconsistent with one's own autonomy, that is, if one's freedom conflicts with itself (see G 4:438, also 421f, 424).

We might do better, however, to construe self-esteem, as the *principium* of duties to oneself, as a subjective principle, and in particular, as a disposition (or attitude) [*Gesinnung*]. This interpretation is suggested by Kant's account of ethics as distinct from right at the beginning of the Collins treatment of *Ethica*. Ethics, Kant says, concerns "the inner goodness of actions" and "refers solely to dispositions [*Gesinnungen*]" (C 27:299; see also 300, 317; cf. V 27:582, 585; MS 6:392f): "In spirit, the moral law ordains the disposition, in its letter the action. We shall therefore see in ethics how the moral law is exercised in spirit, and will not be adverting to the action at all" (C 27:301; cf. KpV 5:152). Kant here describes *Gesinnungen* as "the basic principles [*Grundsätze*] of our actions" which "serve to couple actions with their motivating ground [*Bewegungsgründe*]" (C 27:299, *sic*). Kant's later works characterize a disposition or attitude as a "determining ground" for the agent's adoption of maxims and a "subjective principle of maxims" (KpV 5:125; RGV 6:37). So construed, self-esteem is a subjective practical principle which grounds the choice of, and is displayed in, the actions Kant identifies as duties to oneself – and in other, more specific subjective principles (maxims or rules) underlying these actions.¹⁵ It is a principle of acting in keeping with the dignity of humanity in our own person from esteem for the ideal of humanity in our own person. Taking Kant's claim that self-esteem is the *principium* of duties to oneself to refer to this attitude is congenial to rendering "*principium*" as "origin," "source," or "beginning,"

¹³ Kant presents similar principles elsewhere (see KpV 5:87; V 27:601; MS 6:395, 410). For language in Collins that more closely approximates that of the formula of humanity, see C 27:343.

¹⁴ Kaehler appears more cogent here.

¹⁵ It is not only the moral *attitude* of self-esteem one could construe as a source of duties to oneself; cf. MS 6:403 on the moral *feeling* as their basis.

which sometimes seem like more natural readings of “*principium*” than does “principle.” If we construe “*principium*” in this other way, we can understand self-esteem as the moral attitude from which duties to oneself originate, or in which they begin. I now turn to a more developed account of this attitude.

1.5 Attitude

To be morally acceptable, an attitude of *Selbstschätzung* must at least be consistent with the fundamental moral attitude of obedience to the moral law. My interest, however, is in a more robustly moral form of self-esteem: an attitude we morally ought to have, which pertains to the dignity of humanity in our person (see C 27:347, 349; MS 6:435–437). Such an attitude would, in a person with a morally good character, be firmly grounded in the fundamental moral attitude.

Kant sometimes seems to regard humility in relation to the moral law as an aspect or manifestation of the moral attitude of self-esteem (see, e.g., C 27:348f). Often however, he presents self-esteem in ways that suggest that it is distinct from humility, and humility a necessary complement to it (see C 27:347, 357; V 27:667, 682; MS 6:435, 459; KpV 5:128).¹⁶ As an attitude concerning our dignity as free, moral beings, self-esteem seems more aligned with pride than humility. Indeed, in the *Tugendlehre* Kant glosses “real self-esteem” as “pride in the dignity of humanity in our own person” (MS 6:459). As a moral attitude, being proud of the dignity of humanity in our own person involves not simply being conscious of this dignity, but embracing it as ours to uphold, assert, honor, and realize.

It is largely through pride that moral self-esteem shapes our interactions with others (see C 27:349; V 27:708; MS 6:465).¹⁷ In the *Tugendlehre*, Kant defines “pride proper” (or “noble” pride), as “a concern to yield nothing of one’s own human dignity in comparison with others” (MS 6:465). In *Vigilantius*, Kant states: “Every human being is called upon to concede nothing to another, insofar as the latter seeks to arrogate to himself a precedence over us; this is legitimate pride, which is unwilling to yield anything to the other’s presumptuousness, because we would thereby render ourselves inferior” (V 27:708). Kant contrasts pride, “which is

¹⁶ On Kantian humility, see Grenberg 2005; Loudon 2007a.

¹⁷ Kant often presents noble pride as moral self-esteem manifested in relation to others – in contrast to humility, as moral self-esteem manifested in relation to the moral law (e.g., C 27:349). Limiting noble pride to other-regarding contexts is problematic if we wish to identify it with the virtue of love of honor, as Kant seems to at MS 6:465. Love of honor appears broader; it is manifested in some purely self-regarding contexts, for example, against avarice (MS 6:420, 432).

addressed only to preserving parity of respect for oneself” with arrogance, which is “founded on the demand a person makes, that others should hold themselves in low esteem, in comparison with himself; this, therefore, is associated with a demeaning of the other’s personhood” (V 27:708; see also MS 6:462, 465). Noble pride rebuffs others’ arrogance, refusing to countenance insults to our dignity.

Kant also, in the *Tugendlehre*, identifies noble pride with love of honor, *Ehrliche* (see MS 6:465). The latter notion is slippery, however. Kant’s discussion of it in the lectures is shaped in part by Baumgarten’s discussions of honor, reputation, and *philotimia*.¹⁸ Kant seems to have more than one referent of “*Ehrliche*.” His conception of love of honor appears to evolve along with other aspects of his moral thought, such as his view of moral motivation.

The Collins notes present the love of honor integral to all self-regarding duties as consisting largely in a concern with inner honor and acting in ways that comport with the dignity of humanity in our own person (see C 27:347). Yet here and in other early works, notes, and lectures – before Kant had fully formed his mature view about the feeling of self-esteem as part of the subjective aspect of moral motivation – love of honor is often presented as something originally nonmoral but morally useful: for example, as a drive within human nature that can (and should) be fostered and employed to facilitate moral action.¹⁹ Kant describes love of honor as a gift of nature or providence, distributed universally but unequally among human beings (see C 27:408; also BGSE 2:217f, 227). Love of honor is identified as “an inclination directed solely to honour,” through which providence assures that “no man, even a great one, is indifferent to the opinion of others” (C 27:408; see also 410f). Kant thinks that human beings find contempt “unbearable” (C 27:407) and that self-degradation inevitably elicits contempt. People esteem one another based on moral goodness, and judge moral goodness by appeal to the moral law; so an inclination to honor can serve as a reliable supplement to – and perhaps stand-in for – the pure moral motive. Furthermore, because human beings dread contempt, we can foster aversion to self-regarding vices in children by teaching them to associate them with it.²⁰

The Vigilantius account of love of honor has much in common with that in Collins, though it reflects a more mature stage of Kant’s ethical thought.

¹⁸ See *Ethica*, §§ 293–300, 481–490 (AA 27:815–817, 864–866, 945–947, 1010–1013).

¹⁹ See Denis 2014.

²⁰ See especially Kant’s comments about lying, at M I 27:1429f; ÜP 9:478, 484, 489.

Here a more developed distinctly moral conception of love of honor operates alongside nonmoral conceptions of it, which sometimes serve as its foil (see V 27:666, 680, 695). Kant distinguishes love of honor from the craving for honor (see V 27:666f; C 27:408–410); one who loves honor is modest, and need not have his honorableness acclaimed by others; he “sets out from self-esteem” (V 27:666). Love of honor is “founded upon *true honour*, in contrast to vanity, or all that false honor which does not constitute a person’s worth, e.g., fine clothing” (V 27:664). True honor “rests on the worth that is conferred only through morally good conduct” (V 27:664); one with true honor “deserves to be recognized as worthy of humanity” (V 27:665). In addition to this basic, “negative” sense of true honor, Kant expounds a “positive” sense, which “takes on actions that involve merit” in relation to others (V 27:665) – and, more generally, which go beyond what one’s strict duties require (see V 27:621f, 668). Kant contrasts “that self-esteem which underlies true honour” with “self-abasement, i.e., the disposition whereby we renounce the respect which others may have for us,” which produces “the opposite” of love of honor (V 27:667).

Within the *Tugendlehre*, love of honor, described as a way of thinking [*Denkungsart*], is presented as a virtue (see MS 6:420).²¹ Kant identifies it not only with noble pride in relation to others (MS 6:465), but also, more fundamentally, with just esteem for oneself, “*justum sui aestimium*” (MS 6:420). Although love of honor is manifested externally as respectability, “*honestas externa*” (MS 6:464), its primary concern is inner honorableness, “*honestas interna*” (MS 6:420). Love of honor as a virtue embodies pride in the dignity of humanity in one’s person. Integral to this virtue is a strong commitment to act, in relation to oneself and others, only in ways that comport with that dignity. A person’s character is constituted by her maxims, attitudes, and ways of thinking (see G 4:393; KpV 5:151–153; ApH 7:292f; cf. Kuehn 2009, 15–25). Love of honor and the moral attitude of self-esteem underlying it help constitute a morally good character.

1.6 Proper self-esteem

Myriad relations hold among the morally-grounded forms of *Selbstschätzung* explicated here. Just self-estimation yields a correct opinion of one’s worth and prompts one to feel both humility and elevation.

²¹ See Denis 2014; Bacin 2013. Kant continues to hold that human beings have a natural drive for honor (see MS 6:216).

The moral attitude of self-esteem is displayed in acts necessitated by the objective principle of self-esteem; these acts have the feeling of self-esteem as their proper incentive. A person with the moral attitude of self-esteem values herself in a manner consistent with the just self-estimation.

All moral forms of *Selbstschätzung* concern the dignity of humanity in our own person. Humanity in our person is interchangeable with the moral law as the standard for estimating the worth of the human being in our person. Our opinion of our personal worth must be high with respect to humanity in our person, low with respect to the human being in our person. The feeling of self-esteem is a feeling of the sublimity (dignity²²) of humanity in our person. The objective principle of self-esteem requires action in keeping with the dignity of humanity in our own person. The moral attitude of self-esteem is one of pride in the dignity of humanity in our own person; it is a subjective practical principle of conformity with the dignity of humanity in our own person.

These connections point toward a broad, inclusive notion of proper self-esteem as a morally correct way of valuing ourselves – a multifaceted way of valuing ourselves, which involves the various morally grounded forms of *Selbstschätzung* explicated within sections 1.1–1.5. Consistent with the Collins section “Of Proper Self-Esteem” [*Von der geziemenden Selbstschätzung*], it includes both humility and pride (C 27:348). Indeed, proper self-esteem consists in a web of interrelated and mutually reinforcing beliefs, desires, feelings, and principles concerning our worth as determined by the moral law (see C 27:348–350; V 27:695; MS 6:441; KpV 5:73). Valuing ourselves this way deeply and pervasively shapes our moral outlook, aligning our perspective more closely with that of the moral law.

2 Duties to oneself

In the three works on which we are focused, Kant offers somewhat varied accounts of duties to oneself. I have already discussed some important relations between particular facets of proper self-esteem and duties to oneself. Here I show that within all three accounts, duties to oneself can be seen as expressions of the “morally ordained” attitude of self-esteem.

2.1 Three accounts

In Collins, Kant asserts: “The *principium* of the self-regarding duties . . . consist[s] in . . . self-esteem; our actions . . . must be in keeping with the

²² On dignity and sublimity, see Sensen 2011, esp. 165–168, 172, 178f.

dignity of humanity” (C 27:347, translation modified; see also ÜP 9:488f). Our actions are in keeping with the dignity of humanity when we limit our use of freedom not by inclination, but by agreement with itself. We act in a manner unworthy of humanity in our own person when we undermine the conditions of our greatest use of freedom, such as through drunkenness (C 27:341, 346) or suicide (C 27:342f, 370). To act in keeping with the dignity of humanity in our own person also requires “so disposing [oneself] that [one] may be capable of observing all moral duties” (C 27:348) and self-mastery, whereby “we give morality the supreme power over our selves, so that it rules over our sensibility” (C 27:361).

In *Vigilantius*, Kant explicates duties to oneself in relation to the right of humanity in our own person (see V 27:603). The right of humanity in our own person is its authority to compel the human being in our own person; it underlies not only duties to oneself, but all duties: “To make a rule for oneself presupposes that we set our intelligible self, i.e., humanity in our own person, over against our sensible being, i.e., the human being in our own person, and thus contrast the human being as an agent with humanity as the law-giving party” (V 27:579; see also 509f).²³ Kant depicts perfect duties to oneself as duties to which one is bound most immediately through the right of humanity in our own person (see V 27:601, 603); these duties constrain the human being’s use of his person to conditions of agreement with the right of humanity in his own person. They forbid our disposing over our body as property, giving others unrestricted use of our powers, and robbing ourselves of our honor (V 27:593f). Kant explains the imperfect duty to perfect ourselves thus: “In regard to our own person, humanity is an ideal, to which we owe the duty of perfecting ourselves, so that we may fulfill the duties that it imposes on us” (V 27:626; see also 583).

In the *Tugendlehre*, Kant presents duties to oneself as concerned with the recognition, preservation, and enhancement of our moral agency. Kant associates perfect duties to oneself with “moral self-preservation” and “moral health” (MS 6:419); they preclude action on maxims opposed to inner freedom or conditions of its expression. Kant associates imperfect duties to oneself with “moral prosperity” (MS 6:419); they concern developing our natural talents and moral capacities so as fully to realize our inner freedom. Kant’s analysis of the *concept* of duties to oneself in the *Tugendlehre* recalls the *Vigilantius* analyses of the right of humanity in

²³ In contexts like this, Kant sometimes identifies humanity in our own person with *homo noumenon* (see MS 6:239; V 27:593, 601). On humanity, *homo noumenon*, and *homo phenomenon*, see Wood 2008, 171f; Denis 2010c, 127–131; Byrd and Hruschka 2010, 285–293; and Sensen 2011, 127–133.

our own person. In explaining why the impossibility of duties to oneself would entail the impossibility of all duties, Kant states: “I can recognize that I am under obligation to others only insofar as I at the same time put myself under obligation, since the law by virtue of which I regard myself as being under obligation proceeds in every case from my own practical reason; and in being constrained by my own reason, I am also the one constraining myself” (MS 6:417f).²⁴ Thus here as well as in *Vigilantius*, the inner freedom of moral self-constraint is fundamental to Kant’s conception of duties to oneself.

2.2 Expressing the attitude

Despite the differences among these accounts, in all three, duties to oneself can be fruitfully understood in relation to the moral attitude of self-esteem. Different types of duties to oneself display this attitude in different ways.

Perfect duties to oneself instantiate the attitude of proper self-esteem negatively: through rejection of (e.g.) suicide, self-enslavement, and servility. Collins expounds violations of a wide range of duties to oneself in terms of dishonor, contempt, and worthlessness (see C 27:341f). Flattery is “a want of self-esteem, where we do not scruple to demean our own worth beneath another’s, and elevate his, in order to gain something thereby” (C 27:449f). Everyone “is bound to uphold his rights, and to see that others do not trample them underfoot . . . for otherwise . . . he throws away his humanity” (C 27:435).

In *Vigilantius*, perfect duties to oneself are divided into duties concerning substance, causality, and *commercium* (see V 27:593, 601f). Self-esteem, particularly as it bears on inter-personal relations, is pertinent especially to the third class: “In regard to the *commercium* with others, or the relationship of human beings to one another within society, the agent is prohibited from letting himself be robbed of his honor, or robbing himself of it. Such debasement does not permit the respect he must have for humanity” (V 27:594). One has a strict duty to humanity in one’s own person not only to be honorable, but also to preserve one’s reputation as an honorable human being: “he must retain the respect of others for himself, i.e., his honor, he must in no way give ground through his behavior for any damage to his good name (*bona ex aestimatione*); for example, under suspicion of a crime, he *must* exculpate himself” (V 27:602). These notes also provide a lengthy and diverse list of violations of the “individual rights of humanity in our own person” (V 27:604) (i.e., perfect duties to oneself),

²⁴ See Denis 2010b, 183–186; and Timmermann 2013b.

depicted as directly contrary to moral self-esteem. The violations range from lying and begging through despondency and idolatry (see V 27:604–607; cf. MS 6:436f).

Perfect duties to oneself as a moral being only are the class directly linked with self-esteem in the *Tugendlehre* (see Bacin 2013). Indeed, the human being's duty to himself as a moral being only is, essentially, practically instantiating this attitude through maxims: the duty “consists in what is formal in the consistency of the maxims of his will with the *dignity* of humanity in his person” (MS 6:420, 435f). The vices contrary to this duty directly oppose our inner freedom – our “innate dignity” (MS 6:420) – on which character and virtue depend. Through lying, a human being “makes himself contemptible in his own eyes and violates the dignity of humanity in his own person” (MS 6:429). Miserly avarice is “slavish subjection of oneself to the goods that contribute to happiness” (MS 6:434). In servility, we waive our claim to moral worth in order to attain worth or favor through others, thus disavowing the dignity of humanity in our person (see MS 6:435). Yet Kant's expositions of duties to oneself as an animal and moral being also appeal to self-esteem. These expositions urge us to subject humanity in our person neither to “debasement” [*Abwürdigung*] through suicide, “disposing of oneself as a mere means to some discretionary end” (MS 6:423); nor to “defilement” [*Schändung*] through unnatural lust (MS 6:424), a “complete abandonment of oneself to animal inclination,” which “deprives [a human being] of all respect for himself” (MS 6:425; see Denis 1999, 225–248).

Imperfect duties to oneself express the moral attitude of self-esteem positively: by promoting our moral and natural perfection. We thereby make ourselves as human beings more worthy of humanity (see MS 6:386f, 391–393, 444–447; V 27:626, 651f). In Collins, Kant takes only moral perfection to be a duty: “nothing will come into it, save how much we make ourselves more perfect in regard to our inner worth, and how we are to uphold the dignity of humanity in regard to our own person” (C 27:363, translation modified). Cultivation of self-mastery and self-possession – akin to imperfect duties to oneself in certain respects, though not always classified as such – strengthens the inner freedom integral to human moral agency and virtue, and is thus essential to maintaining and realizing the innate dignity of humanity in our own person (see C 27:360–369; V 27:626f, 654; MS 6:407–409).

All three accounts of duties to oneself present love of honor as integral to duties to oneself – and as an especially important duty to oneself. Collins says that love of honor is “an object of our duty. The human being must be

honor loving” (C 27:411). Kant here describes love of honor as the ground, source, and essential ingredient of duties to oneself (see C 27:347, 411f). Vigilantius describes love of honor as “the highest duty of humanity to oneself, so little capable of abridgment, that it has to go further than love of life” (V 27:664; see also C 27:373, 378). There are conditions under which a human being is not bound to preserve her life; there are no conditions under which she is not bound to uphold the dignity of humanity in her own person (V 27:665). The *Tugendlehre* presents love of honor as the virtue which correlates directly with the duty to oneself as a moral being only and opposes lying, avarice, and servility – vices directly antagonistic to the moral attitude of self-esteem, as well as to inner freedom, moral character, and virtue (MS 6:420).

3 Primacy, sanctity, and worthlessness

The [previous section](#) not only highlights tight connections between duties to oneself and the moral attitude of self-esteem but also suggests why Kant attributes primacy to duties to oneself. On all three accounts, duties to oneself are immediately grounded in, and essentially concerned with, the freedom in which morality consists (see Denis [2010b](#)). It is no mere coincidence that all three accounts also portray duties to oneself as grounded in a principle of respect for humanity in one’s own person. Humanity is the free, moral, intelligible being within us, or the human being conceived simply in terms of its freedom (KpV 5:87; MS 6:239; V 27:543, 579, 593). Human beings have our elevated standing – our innate dignity – in virtue of the humanity (freedom) in our own person.

Although we can explain the primacy of duties to oneself without referring explicitly to their relation to proper self-esteem as a whole or to the moral attitude of self-esteem specifically, attending to these relations enhances our grasp of their primacy. Acting in keeping with the dignity of humanity in our own person – that is, consistently with one’s autonomy – is the objective condition of morality. Just self-estimation, and the feelings of humility and elevation resulting from it, reinforce a proper orientation of the agent to the moral law, and (derivatively) to other human beings. The moral attitude of self-esteem is essential to our moral self-preservation and moral health as human agents. Pride in the dignity of humanity in our own person is crucial to sustained pursuit of our moral vocation, and thus to the fulfillment of duties to others as well as to oneself. Love of honor combats vices incompatible with character. Duties to oneself not only manifest but also promote proper self-esteem, which is a significant if diffuse aspect of the moral life.

Furthermore, attention to relations between proper self-esteem and duties to oneself helps us understand why Kant talks as he does about *violations* of duties to oneself. An aspect of the primacy of duties to oneself is (what I call) their sanctity: their violation constitutes a distinctive and extreme sort of wrong. I suggest that it is as attempts to convey with maximal force the sanctity of duties to oneself that we should understand Kant's portrayals of the violations of these duties as rendering the agent worthless, contemptible, and undeserving of respect. Kant wants his audience to *feel* the sanctity of these duties. The preceding sections put us in a good position to appreciate Kant's strategy – appealing to our feeling for our moral vocation, desire for inner worth, and dread of contempt – as both philosophically justified and psychologically compelling.

In Collins, Kant supports his claim that duties to oneself are “the most important of all” by saying: “He who violates duties toward himself, throws away his humanity” (C 27:341); “the infringement of self-regarding duties takes all his worth from a human being” (C 27:341). He then “illustrate[s] the violation of self-regarding duties with a few examples,” all of which stress the ways in which such violations bring the human being into contempt, degrade humanity in his person, and so on (C 27:341–343). Similarly, in *Vigilantius*, Kant says that by violating perfect duties to oneself: “[W]e thereby make ourselves unworthy of the possession of our person that is entrusted to us, and become worthless, since the preservation of our own worth consists solely in observing the rights of our humanity. We lose all inner worth, and can at most be regarded as an instrument of others, whose chattel we have become” (V 27:604). To impress upon his audience that perfect duties to oneself are “the highest duties of all,” Kant enumerates violations of them, emphasizing the degradation they involve and the contempt they draw (V 27:604–606).

Perhaps the most disturbing claims within such passages suggest that by violating a duty to oneself, a human being forfeits his standing as a person, and may be treated by others however they wish: “he can be treated by others as an animal or a thing, he can be dealt with like a horse or a dog, for he is no longer a human being” (C 27:373, see also 346). Dramatic statements along these lines are more common in Collins than in *Vigilantius*. In the *Tugendlehre*, Kant makes a relatively mild statement of this sort: “one who makes himself a worm cannot complain afterwards if people step on him” (MS 6:437; cf. 420, 429).

A great deal is going on in passages in which Kant depicts the violation of duties to oneself as entailing the worthlessness of the agent. The main idea, however, seems to be this. In violating duties to oneself, a duty to which

humanity in one's own person immediately binds one, one directly disregards the authority and dignity of humanity in one's own person, and so also one's status as a person and the condition of attaining inner worth through virtue. To the extent one can make oneself worthless, one does so by violating duties to oneself. As we have practically denied our status as persons, we are in no position to demand that others respect us as such (see M I 27:1427f; C 27:373). Finally, by disavowing our moral agency, the dignity that goes with it, and the inner worth it allows us to attain, we predictably elicit contempt, which implies a judgment of worthlessness (see V 27:709; MS 6:463).

Kant's dramatic claims about agents' making themselves dogs, chattel, things, etc. by violating duties to oneself should *not* be understood to imply that it is morally permissible to treat arbitrarily someone who has violated such a duty. One does not rid oneself of the binding authority of humanity in one's own person by flagrantly violating duties to oneself. One remains morally obligated to and through humanity in one's own person – indeed, obligated to make oneself worthy of it. Correlatively, humanity in the person of another human being is something our own reason compels us to recognize practically even if she does not. Humanity in her person limits how we may treat even the basest human being, who cannot consistently demand respect for herself from others (MS 6:463f).²⁵

Kant makes these provocative claims not to tell us how we may treat *other people* who violate duties to oneself, but rather to emphasize how utterly contrary to the dignity of *humanity in our own person* violations of such duties are and to incite our aversion to them. By talking about the sanctity of duties to oneself in terms of the worthlessness implied by their violation and the contempt thus elicited, Kant makes vivid – nearer to intuition and feeling, one might say (see G 4:437) – the fundamental importance of complying with them. This interpretation is supported not only by the use of this language explicitly to impress upon his audience the significance of these duties (see C 27: 341–343; V 27:604–606), but also by Kant's saying, in Collins, that consideration of the consequences of violating our self-regarding duties helps “to provide a better insight into the[ir] *principium*”

²⁵ Occasional comments suggest that only someone with no respect for humanity in her person, no interest in living up to the standard set by the moral law, or no pride in her moral vocation, could act contrary to her duty to herself (see C 27:341). Given that the feeling of self-esteem is a subjective condition of human morality (see MS 6:402f), violations of strict duties to oneself might suggest that the human being who violates them is not a moral agent. Such a being, however, could have no duties to violate. So if we regard beings as violating duties to themselves – rather than as amoral beings who cannot be blamed for (e.g.) maiming themselves – we see them as persons, and cannot coherently treat them as things.

(self-esteem) and “the inner vileness” of their violation (C 27:347). Given his views that human beings naturally desire respect and dread contempt (see C 27:407f), and that not even the worst of us is beyond wishing for moral worth (see C 27:418; G 4:454; KpV 5:87f), Kant would regard referring to worthlessness and contempt as a compelling way to convey the sanctity of duties to oneself.

The preceding analysis helps explain Kant’s dramatic claims concerning violations of duties to oneself. It does not, however, address why Kant makes them with varying frequency, flair, and intensity among Collins, Vigilantius, and the *Tugendlehre*. I conjecture that there are rhetorical, and perhaps pedagogical, reasons as well as philosophical ones. It is unsurprising that this approach is more prevalent, prominent, and extreme in the lectures than in the *Tugendlehre*. A lecture calls for greater flamboyance than does a published philosophical text. Moreover, Kant may have regarded his students as a more appropriate target than his readership for such blatant, provocative appeals to honor- and worth-related feelings and desires – appeals that might increase their aversion to violations of duties to oneself and perhaps foster in them moral forms of self-esteem. That the Collins notes have more pervasive claims of this kind – and more numerous of the more severe and disturbing sort (e.g., C 27:346, 373; cf. V 27:604) – than the Vigilantius notes is consistent with the greater role that self-esteem and love of honor play generally in the Collins account of the grounding and nature of duties to oneself.²⁶

²⁶ I thank Stefano Bacin, Oliver Sensen, and Mark Timmons for correspondence regarding “*Selbstschätzung*,” and Stefano Bacin, Oliver Sensen, and Eric Wilson for comments on various drafts of this essay.

*Virtue, self-mastery, and the autocracy
of practical reason*

Anne Margaret Baxley

Kant's theory of virtue has received considerably less attention than the moral theory set out in his most widely read ethical texts, the *Groundwork of the Metaphysics of Morals* and the *Critique of Practical Reason*. Moreover, this account of virtue has not always been well understood. In light of the fact that Kant characterizes virtue in terms of strength of will over feelings and inclinations that conflict with duty, commentators have suggested that Kant's account of virtue is impoverished, insofar as it appears to be a recipe for nothing more than continence.¹ Part of the problem in uncovering a comprehensive picture of Kant's full account of moral character lies with Kant, for his remarks about virtue in his later and less well-known writings are scattered and not always systematically presented. Although the *Doctrine of Virtue* contains a sustained treatment of virtue and our various ethical obligations (duties of virtue), that work requires careful interpretation and reconstruction. Fortunately, Kant's *Lectures on Ethics*, now easily accessible in English translation, provide an additional valuable resource for readers interested in understanding a more complete picture of Kant's conception of virtue, its importance in his overall ethics, and its relation to other foundational concepts in his moral theory.

This essay analyzes four central theses concerning the nature of virtue that Kant himself emphasizes in the *Lectures on Ethics*. Although these four themes do not provide an exhaustive account of Kant's theory of virtue, together they lay a solid foundation for any systematic interpretation of Kant's considered views on moral character.

¹ This distinction between full virtue and mere continence, which is crucial for Aristotle's moral psychology, is understood to be a distinguishing feature of classical virtue ethics more generally. Contemporary theorists who have questioned whether Kant has the resources to distinguish between the person who merely acts rightly and the person who is wholehearted in what she does include Julia Annas, Talbot Brewer, David Brink, Rosalind Hursthouse, and Martha Nussbaum. See Annas 1993, 53 and 2006, 517; Brewer 2002; Brink 1999, 580 and 2000, 431; Hursthouse 1999, 104; and Nussbaum 2001, 172. I explore this theme at length in Baxley 2010.

I Virtue involves moral self-mastery and moral strength of will

In the *Lectures on Ethics*, Kant appeals to the key notion of self-command or self-mastery (*Selbstbeherrschung*) to characterize virtue. Self-control in the service of prudence, which ultimately aims at the satisfaction of inclination, is not sufficient for virtue, for, as Kant puts it, “true self-control is moral in character” (C 27:361). The self-control crucial for moral agency is distinguished from mere prudential self-control in that it involves mastering and ruling the rabble of sensibility so that one possesses the fortitude to do one’s duty from duty in spite of any obstacles that might stand in the way (C 27:360–362). Virtue, as Kant thus conceives of it, primarily involves power, strength and authority (C 27:465; V 27:492). More specifically, it is a distinctly *moral* strength of will or soul, where strength connotes the force to withstand all temptations to transgress the moral law and is manifest in a victory over inclination (C 27:465; M II 29:603–606; V 27:491f).²

In his discussions of virtue, Kant draws a distinction between an executive and a directing power of the will, indicating that self-mastery is distinguished not only from mere prudential self-regulation, but also from the more familiar Kantian conception of self-legislation. In his words: “We have a double authority over ourselves, the disciplinary and the productive. The executive authority can compel us, in spite of all impediments, to produce certain effects, and in that case it has might. But the directing authority exists merely to guide the forces of the mind” (C 27:363). The notion of a disciplinary or executive authority that Kant invokes here in an effort to highlight what is characteristic of self-mastery is described as an acquired strength of will to compel ourselves to act in accordance with norms of pure practical reason (norms that we give ourselves in virtue of our directing authority). The thought behind this distinction is that, as self-legislating beings with autonomous wills, we are capable of directing ourselves to comply with moral laws grounded in pure practical reason alone. More than this legislative authority is required of us, however, if we are to live fully in accord with the demands of pure practical reason, executing the moral law. For this task, we need self-mastery, the actual moral strength of will to weaken a strong (internal) opponent to duty and to put the law into practice. In underscoring the idea that such rational self-command is the very condition under which we finite rational beings have the power to resist being determined to act by

² For Kant’s account of virtue as moral strength of will in the *Metaphysics of Morals*, see MS 6:389, 393, 397, 405, and 408f.

strong natural inclinations and to conform our wills to self-legislated moral commands, Kant explains:

That man, moreover, should act in accordance or adequacy with the moral laws, can occur insofar as he has repressed and conquered, through the moral law, the inclination he harbors to deviate or do the opposite. The struggle of inclination with the moral law, and the constant disposition (*intentio constans*) to carry out his duties, therefore constitutes what we call *virtue*. (V 27:492)

It should come as no surprise that being one's own master in this Kantian sense requires self-discipline and self-compulsion (*Selbstzwang*) (M II 29:616–618; V 27:520f). In disciplining herself, the virtuous agent plays the role of sovereign over herself, or, better, is one in whom morality itself has sovereign authority (C 27:361). In light of the fact that passions, as Kant conceives of them, are pernicious forms of inclination that conflict with moral concerns and lead readily to vice, moral self-discipline involves ruling oneself so that one has no passions.³ In short, morality commands us to yield nothing to passion, a point that leads Kant to praise the Stoics for recommending the extirpation of passions as the proper path to virtue (C 27:368).⁴

Of course, the notion of self-compulsion might sound like a harsh or repressive recommendation for acquiring moral character. Kant, however, believes that self-discipline carries with it its own rewards. As a form of strength, virtue comes in degrees, and the more disciplined and hence virtuous a person is, the more he is free (V 27:520).⁵ In emphasizing this close connection between virtue, self-compulsion, and freedom, Kant insists:

The more a man considers a moral act to be irresistible, and the more he is compelled to do it by duty, the freer he is. For in that case he is employing the power he has, to rule over his strong inclinations. So freedom is all the more displayed, the greater the moral compulsion. (M II 29:617)

³ Kant holds that affects (*Affekte*) and passions (*Leidenschaften*) are particular forms of feelings and inclinations that are inherently problematic from the standpoint of both morality and prudence, because they are directly contrary to reason. For Kant's remarks about affects and passions in relation to virtue and self-mastery in the *Lectures on Ethics*, see C 27:368 and V 27:612. Kant's official account of affects and passions in his published writings can be found in MS 6:407–409; ApH 7:251–282; KU 5:272n.; and RGV 6:29n. For a helpful discussion of the connection between Kant's views regarding affects and passions and the Stoics, see Sherman 1997, 121–186. Julia Annas' analysis of the Stoics' account of emotions also includes a useful comparison of the Stoics and Kant. See Annas 1993, 53–66.

⁴ Although Kant shares the Stoic view that passions are wholly at odds with reason and must be avoided, he expresses doubts about how far we can progress in getting rid of them (C 27:368).

⁵ For an excellent discussion of Kant's conception of virtue as a source of inner freedom, see Engstrom 2002.

Moreover, the self-commanding agent enjoys a sense of equanimity that stems from the fact that he does not experience strong inner turmoil between duty and contrary inclinations.⁶ The state of the soul of the agent lacking in virtue, by comparison, is characterized in terms of ongoing conflict and strife. As Kant characterizes the difference: “Without disciplining his inclinations, man can attain to nothing, and hence in self-mastery there lies an immediate worth, for to be master over oneself demonstrates an independence of all things. Where there is no such self-mastery there is anarchy” (C 27:361f). Still worse, the vicious person is portrayed as being enslaved by the power of natural inclinations (C 27:464). In light of the fact that moral self-discipline and self-mastery promotes freedom from determination by strong inclinations and signifies a condition of the soul akin to psychic harmony (a sense of inner freedom), it is easy to see why Kant holds that virtue is its own reward.⁷

2 Autocracy is the ideal form of moral self-governance required for virtue

Kant’s various comments concerning the nature of virtue and self-command in the *Lectures on Ethics* indicate that he belongs to a rich philosophical tradition of theorizing about virtue that relies heavily on political metaphors about governance to illuminate moral ideals about self-governance. We are told that the self-mastery required for virtue provides a model for a “vigilant government,” one in which reason must exercise force in relation to sensibility in order to “compel this rabble under the rule in accordance with ordinance and regulation” (C 27:360). The specific term Kant uses to characterize his favored ideal form of moral self-governance is *autocracy*. Although he explicitly appeals to the term only once in his most extended published treatment of virtue, the *Doctrine of Virtue*, it appears or is suggested more frequently in his remarks about virtue and self-mastery in the *Lectures on Ethics*. This notion of autocracy underlies

⁶ In the *Critique of Practical Reason*, Kant claims that virtues involves consciousness of mastery over inclinations and produces a sense of independence from inclinations, thereby affecting a feeling of moral contentment that Kant describes as a negative form of happiness (KpV 5:118f). For other passages in which he elaborates on the relation between virtue, self-compulsion, freedom, and moral contentment, see M II 29:617; V 27:520, 643, 647f; KpV 5:38, 111–119, 155f, 161; and MS 6:380, 406f.

⁷ Kant repeatedly warns against recommending virtue for its beneficial consequences, insisting that virtue should be praised because it possesses its own inner worth and is its own reward. For Kant’s account of virtue as its own reward, see G 4:426n., 442f, 450; and MS 6:391, 396, 406.

Kant's considered account of moral character, providing the key to understanding his distinctive concept of virtue.⁸

Kant introduces the notion of autocracy by insisting that true moral strength of will involves autocracy in addition to autonomy, noting that, "[i]f reason determines the will through the moral law, it has the force of an incentive, and in that case has, not autonomy merely, but also autocracy. It then has both legislative and executive power" (M II 29:626). Autocracy signifies the self-command at the heart of virtue, for it represents "the authority to compel the mind, despite all the impediments to doing so. It involves mastery over oneself, and not merely the power to direct" (C 27:362). As actual strength of will to resist forces in conflict with duty, autocracy is consistently portrayed as an executive authority or power of the will, which Kant contrasts with a separate legislative authority or power to guide the will in accordance with norms of pure practical reason (C 27:362; M II 29:626). Kant speaks of an autocracy of "the moral law" (or "laws"), "reason," or "freedom," but also refers to an autocracy of the "soul," "heart," or "mind" over inclinations.⁹ When he appeals to the concept of autocracy in the *Doctrine of Virtue*, he uses it in precisely the same way as in the *Lectures on Ethics*, namely, to underscore its importance for his overall picture of virtue as acquired moral strength of will and to distinguish it from autonomy. Having differentiated between a doctrine of morals, which he explicitly connects with autonomy, and a doctrine of virtue, which he says "also" involves autocracy, Kant explains that autocracy is distinct from autonomy insofar as it involves "consciousness of the *capacity* to master one's inclinations when they rebel against the law, a capacity which, though not directly perceived, is yet rightly inferred from the moral categorical imperative" (MS 6:383, 515).¹⁰

Kant understands autocracy foremost as a form of self-rule in accordance with the commands of practical reason over sensibility, a model of self-governance presupposing the power to control and limit the influence on the will of feelings and inclinations at odds with moral demands. It stands

⁸ For my extended treatment of Kant's account of autocracy, see Baxley 2010. In addition, see König 1994. For a nice discussion of the intrinsic value of autocracy in Kant's ethics, see Guyer 1993, 346–350 as well as Guyer's later discussion in Guyer 2005, 136–141. Other brief accounts of autocracy can be found in Allison 1990, 164 and Carnois 1987, 120.

⁹ Kant refers to an autocracy of practical reason (MS 6:383) or reason (M II 29:626), an autocracy of moral laws (VMS 23:396), and an autocracy of freedom in relation to happiness (R #6867, 19:186). For passages focusing on autocracy as a power of the heart, soul, or mind, see C 27:362, 364, 368, 378.

¹⁰ Similarly, in his notes to the *Doctrine of Virtue*, he writes that the virtuous person is conscious of "an autocracy (not merely autonomy) of moral laws against all conflicting impulses of sensibility (inclinations)" (VMS 23:396).

for the executive power of the soul to enforce the morally good choices legislated by the directing or guiding power of the will (a power belonging to the rational will in virtue of the property of autonomy) (M II 29:626). The fact that autocracy represents the executive authority of reason's prescriptions in accordance with moral laws underlies Kant's repeated remark that autocracy is the subjective condition of the performance of duty, for, as he puts it, maintaining command over oneself is the condition under which one is able to perform the self-regarding duties and thereby all other duties (C 27:360, 364, 368).

Although Kant tends to identify autocracy with rational rule over sensibility, it is important to note that he conceives of autocracy as a broad and encompassing form of moral self-governance, one that is not exhausted by the simple notion of strength in opposition to sensible feelings and inclinations. Kant holds that human beings possess various mental powers that bear on morality, and he explicitly discusses the ways in which autocracy shapes and transforms these morally relevant powers in the virtuous agent (C 27:364–368). For instance, the autocratic agent is skilled in controlling her imagination, so that her imagination does not tempt her to pursue objects that are beyond her reach or especially hard to attain. She is able to suspend judgment on important matters and is thereby able to refrain from making rash decisions or acting prior to rational reflection. She possesses what Kant calls the “power of presence of mind,” characterized as a harmony of one's mental powers, where such harmony eases the burden of carrying out one's work (C 27:366). In addition, the autocratic agent demonstrates a degree of self-knowledge and capacity for self-examination that the weak-willed agent and the vicious agent lack. By disciplining herself, instead of yielding to emotions and inclinations, the autocratic agent has the fortitude to resist feelings and inclinations that can pose temptations to duty; she displays constancy of mind and the cheerful heart of virtue; and she enjoys a feeling of inner worth that is consequential upon having a firm disposition to do one's duty from duty (C 27:361f; M II 29:617, 623; V 27:520, 643, 648f).

Given Kant's more familiar concept of autonomy, the question naturally arises: what is the precise relation between autonomy and autocracy? This interpretive puzzle is best addressed by recalling the main features of Kant's critical view (first laid out in the *Groundwork of the Metaphysics of Morals*) that autonomy is a property of the rational will and the foundational principle of morality. In the *Groundwork*, Kant defines autonomy as “the property of the will by which it is a law to itself (independently of any property of the objects of volition)” (G 4:440). Autonomy, so conceived,

is essentially a capacity for self-legislation, a capacity implying the freedom to make universal law through one's own will, or freedom to adopt maxims of action that are valid for oneself only insofar they are valid for all other rational agents. According to Kant's particular conception of autonomy, the laws that we give ourselves are prescriptions of our own reason, through which we freely constrain ourselves in light of our recognition of their validity for all rational agents.

On the regressive, analytic method of argument he employs in the *Groundwork*, Kant assumes that morality requires acting on the basis of the categorical imperative. He then introduces autonomy as the supreme principle of morality in the sense of being the necessary condition of the very possibility of morality. The key point to Kant's argument in the *Groundwork* is that acting on the basis of the categorical imperative presupposes a capacity to determine oneself to act independently of, and even contrary to, one's particular interests as a sensuous being with needs, namely, one's empirical interests. Kant treats this notion of self-determination as one that is built into the characterization of autonomy as a property of the will. That is, a will and only a will with the property of autonomy is capable of acting on the categorical imperative, because only a will capable of determining itself independently of its needs as a sensuous being can act in accordance with a practical principle that commands unconditionally because of its mere form.¹¹

This reminder of how Kant conceives of autonomy indicates that the relation between autonomy and autocracy can be captured by distinguishing between a *legislative authority or power of the will* for creating and giving to oneself laws that are universally valid and an *executive authority or power of the will* for enforcing and enacting these laws. The former capacity for self-determination by pure practical reason is one that we all possess in virtue of our autonomy, and it defines our moral personality. The latter power of self-rule by pure practical reason is one that we can acquire through a process of moral self-discipline and self-mastery, and it describes our empirically acquired moral character.

One final point regarding the nature of autocracy is worth mentioning. In spite of the unfortunate political connotations associated with the term, in the *Lectures on Ethics*, Kant denies that autocracy amounts to wholesale self-repression. Whereas Baumgarten endorses the idea that virtue entails self-overcoming or self-conquest (*Selbstüberwindung*) in *Philosophical Ethics* (one of the two texts Kant regularly included for discussion in his course on

¹¹ My understanding of Kant's conception of autonomy is indebted to Allison 1990, 85–106.

ethics), Kant rejects the idea that proper self-governance by reason involves waging a war against oneself. As he explains:

Our author goes on to speak of self-conquest. But if a man rules himself so well that he prevents any rebellion of the rabble in his soul and keeps peace within it (which here, however, is not contentment with everything, but good command and unity within the soul), and if he now conducts so good a government within himself, then no war will arise in him, and where there is no war, no conquest is necessary either. It is therefore far better if a man is so governed that he need gain no victory over himself. (C 27:368f)

In short, although autocracy represents a form of self-governance in which practical reason is sovereign, Kant imagines autocracy to be distinct from both anarchy (where the mob of sensibility rules over reason) *as well as* tyranny (where reason exercises excessive force over sensibility). Kant explicitly cautions us not to go too far in disciplining our natural inclinations. He thinks excessive self-discipline is actually opposed to our duties to self, for we have an obligation to provide for ourselves so that we find satisfaction in life (MS 6:452). It is this very thought – that morality does not require self-denial – that leads Kant to criticize Diogenes for taking things too far and overdoing the duty to do without (V 27:650).¹² Accordingly, Kant rejects the view he labels *ethica morosa* (killjoy ethic) on the grounds that it involves the sacrifice of “all wished-for enjoyment of happiness” (V 27:663). The Kantian autocratic agent, in whom reason reigns supreme, need not conquer herself, but must merely work to maintain constant vigilance over her inclinations so that she knows how to determine herself to act on moral principles in good or ill fortune (V 27:663).

3 Virtue presupposes conflict and struggle and is thereby distinct from holiness

In analyzing the strength required for moral character, Kant regularly contrasts virtue and holiness, largely to underscore the fact that virtue presupposes some element of conflict and struggle. Holiness signifies complete purity of will, even of thought. It represents a state of mind from which an evil desire never arises (M II 29:604). In light of the fact that a holy

¹² Diogenes allegedly lived in a wine jar, possessed only one cloak, and discarded his cup so that he could drink out of his hands, all in the name of simplicity. From Kant’s perspective, however, Diogenes’ life exemplifies the kind of “morose” or “monkish” virtue we should reject, for there is no purpose in punishing ourselves by denying ourselves what we can easily acquire (C 27:393f).

will possesses a pure moral disposition *innately* and consequently wills only what the moral law prescribes, a holy will has no need for virtue, as “the strength of the resolve to perform our duties, and to strive against the constant enticements to do otherwise which sensory feelings inspire” (V 27:570).¹³ Virtue is thus relevant only for finite (sensuous) rational beings, imperfect beings who have needs and inclinations that can conflict with duty and who must exercise self-constraint in order to do what the law commands (M II 29:605f). In emphasizing the notion that virtue entails the real possibility of acting contrary to the moral law and is thus limited in scope to finite creatures in whom there are genuine obstacles to morally good conduct, Kant explains:

Ethics explained by a *doctrine of virtue* is good inasmuch as virtue belongs solely before the inner tribunal; but since virtue entails, not just *morally good* actions, but at the same time a great possibility of the opposite, and thus incorporates an inner struggle, this is therefore too narrow a concept, since we can also ascribe *ethics*, but not virtue (properly speaking) to the angels and to God; for in them there is assuredly holiness but not virtue. (H 27:13)

Readers of Kant's ethics will recall that at times Kant portrays inclinations themselves as the fundamental problem with which we must contend in our efforts to fulfill our moral obligations and live in conformity with the moral law. Inclinations are alleged to be blind and servile and always burdensome to a rational being (KpV 5:118). In both the second *Critique* and the *Groundwork*, Kant insists that inclinations are so problematic that it must be the universal wish of all rational beings to be completely free of them (KpV 5:118; G 4:428). The thought that inclinations are the main impediment to the conformity of the finite rational will to moral laws is likewise implied in passages from the *Lectures on Ethics* in which Kant warns that virtue requires constant vigilance over inclinations and is manifest in a victory over them (V 27:663; C 27:465).¹⁴ Yet, the true opponent of virtue, as Kant sees it, lies much deeper than our inclinations, namely, in an actual tendency to evil on our part, which Kant understands as a tendency we have to subordinate duty to inclination, thereby granting priority to

¹³ For passages in the *Lectures on Ethics* in which Kant contrasts virtue and holiness, see H 27:13; C 27:463f; M II 29:604–606, 611; and V 27:486–492, 518–520. In the *Critique of Practical Reason*, Kant defines virtue as a “moral disposition *in conflict* (*im Kampfe*), and not *holiness* in the supposed *possession* of a complete *purity* of dispositions of the will” (KpV 5:84). For additional passages in his published writings in which he distinguishes virtue from holiness, see G 4:439; KpV 5:32f, 122, 128; and MS 6:383, 396f, 405.

¹⁴ Similarly, in the *Doctrine of Virtue*, Kant characterizes virtue as an acquired capacity “to overcome all opposing sensible impulses” (MS 6:397).

considerations based on our happiness over moral concerns. Virtue is “the strength of soul to withstand, out of duty, the onset of evil,” or “an aptitude, on moral principles, for overcoming the inclination to evil” (M II 29:603; C 27:463). In the *Lectures on Ethics*, Kant characterizes this inclination to evil as an inclination to deviate from the moral law and act in opposition to it (V 27:492). In a notable passage in which he makes explicit the idea that the internal struggle we face in our efforts to conform our actions to the law is not one of merely controlling our inclinations, but overcoming a deep-seated tendency to transgress the moral law, Kant claims:

It is just because he possesses by nature a nevertheless conquerable tendency and propensity to evil, so that he harbors the possibility of being easily drawn away from good into transgression, that duty is coupled, in his case, with moral necessitation, and the latter needed for virtue. (V 27:571)

In *Religion within the Boundaries of Mere Reason*, Kant labels this tendency to evil “radical evil” in the context of arguing for his official doctrine that there is a universal, freely chosen propensity (*Hang*) to evil in human nature.¹⁵ In calling evil “radical,” Kant does not mean to imply that humanity is diabolically evil. Instead, he explicitly states that the term “radical” should be understood in its etymological sense of “roots.”¹⁶ Radical evil for Kant thus signifies the very ground (or root) of the possibility of moral evil in general and of all actions that are contrary to duty (RGV 6:32).

In *Religion*, Kant expounds on this doctrine of radical evil by noting, “the statement, ‘The human being is *evil*,’ cannot mean anything else than that he is conscious of the moral law and yet has incorporated into his maxim the (occasional) deviation from it” (RGV 6:32). In other words, the claim that there is universal propensity to evil in human nature amounts to the claim that we have a tendency to exercise our free choice (*Willkür*) by adopting a fundamental maxim to subordinate the moral law to nonmoral considerations based on happiness and its principle of self-love (RGV 6:36). Simply put, Kant holds that we human beings have a propensity to treat happiness as our supreme or unconditioned good, or a tendency to act in accordance with the dictates of morality only on the condition that what morality

¹⁵ Kant explicitly introduces the doctrine of radical evil in *Religion* (1793), but the outlines of this doctrine can be found earlier, in the *Groundwork* (1785), where Kant writes of “a natural dialectic” of practical reason, understood as a “propensity to rationalize against those strict laws of duty and to cast doubt upon their validity, or at least upon their purity and strictness, and, where possible, to make them better suited to our wishes and inclinations, that is, to corrupt them at their basis and to destroy all their dignity” (G 4:405; cf. G 4:424).

¹⁶ The Latin *radic-* means “roots.” Kant explains that the kind of evil he has in mind is one that is “entertwined [*verwebt*] with humanity itself and, as it were, rooted [*gewurzelt*] in it” (RGV 6:32).

requires of us does not conflict with the satisfaction of our inclinations (the sum total of which is happiness).

Of course, the thesis that humanity is evil in this sense is a substantive thesis about human nature, one that sharply separates Kant from classical virtue theorists. In a discussion in the *Lectures on Ethics* in which he raises this very point, Kant indicates that the various ideals of human excellence found in antiquity fail to provide an appropriate measure of *human* virtue, precisely for the reason that the ancients did not see that finite imperfect beings always require moral strength of will in the face of a powerful (internal) opponent to duty in the struggle to lead morally good lives. From this Kantian perspective, then, Diogenes' perfect man of nature, who lives fully in accordance with reason without struggle, "has no need of virtue, for he has no concept of evil" (M II 29:603). In short, as Kant sees it, a concept of virtue fitting for us mere human beings must retain some element of what the ancients would view as continence – strength of will in the face of opposition to duty – because virtue itself is a corrective (more specifically, a corrective to a strong tendency we have to grant priority to inclinations over the moral law).¹⁷ As Kant thus warns us in the context of reacting to these ancient models of human excellence: "We can attain to virtue, i. e. a moral preparedness to withstand all temptations to evil, so far as they arise from inclinations. The ancient philosophers never got to that point" (M II 29:604).

The key idea that the very concept of virtue presupposes a notion of inner conflict does not imply that the Kantian virtuous agent constantly struggles with actual contrary-to-duty inclinations, always doing her duty from duty in the face of temptation, or that Kant simply identifies virtue with mere continence. On the contrary, the self-commanding agent who has acquired the moral strength of will to govern her sensible nature and successfully hold in check her tendency to evil does not experience the internal turmoil that the continent person and the weak-willed person feel. The Kantian autocratic rules herself in accordance with practical reason and plays the master over herself; as a result, she possesses a form of independence from inclinations and the temptations associated with them. If we recall, the reason Kant rejects the idea that virtue involves winning a battle over oneself (self-conquest) is that autocracy signifies a soul that is unified and at peace,

¹⁷ In the *Lectures on Ethics*, Kant focuses primarily on four particular classical ideals of human excellence: the Cynic ideal of Diogenes and Antisthenes (natural simplicity), the Epicurean ideal (the man of the world), the Stoic ideal of Zeno (the sage), and the Christian ideal of holiness (M II 29:603–605; V 27:647–650). In each case, Kant thinks these ideals miss something at the very heart of virtue, insofar as virtue involves strength in overcoming a serious "hindrance in the will" (a tendency to evil).

and not in a perpetual state of discord (C 27:368f). Still, the point remains that, for Kant, acquiring a virtuous disposition always takes place in a particular context, namely, one in which we must first break with or overcome a deep-seated tendency to give priority to inclination over duty, a tendency we must subsequently remain on guard against, since it can be overcome and repressed and hence not acted on, but never entirely eliminated (C 27:463f).¹⁸

4 **Virtue amounts to a disposition to do one's duty from duty and therefore entails constraint and necessitation by practical reason**

Should there be any doubt as to whether Kant treats the concept of duty as central in his theory of virtue, it is important to see that Kant conceives of virtue as a settled disposition to do one's duty from duty, or out of respect for the law (C 27:308; M II 29:611, 624; V 27:623f, 715). Virtue has its own motive, which springs from the inner goodness of actions (C 27:308).¹⁹ For one's action to have moral merit, it is required that "the action should have arisen from a purely moral intent, or that only the moral law, and not inclination of any kind, should have been the motive for it" (V 27:611). The thought that a morally good disposition involves strength of resolve to fulfill one's duties from the motive of duty, not inclination, helps explain why virtue does not make us happy. In Kant's words:

Virtue does not flit or curry favor, but is honorable. Duty is not what I do for my advantage, but what I do for the sake of the law. The cast of mind which is won over only by reward, is called *indoles servilis*; that which acts only for

¹⁸ The fact that virtue presupposes conflict and struggle with a formidable (internal) opponent explains why Kant describes the process of acquiring a virtuous disposition in terms of a radical "change of heart" or "conversion" (RGV 6:47f; C 27:464). Becoming morally good involves reorienting one's way of thinking or cast of mind (*Denkungsart*), by conquering a tendency to evil and adopting an entirely different fundamental maxim always to subordinate inclination to duty. Moreover, sustaining one's virtuous disposition requires continual effort, for the propensity to evil is a powerful foe that can be overcome but never entirely vanquished. As Kant describes this ongoing ethical task: "However virtuous a man may be, there are tendencies to evil in him, and he must constantly contend against them. He must guard against the moral self-conceit of thinking himself morally good, and having a favorable opinion of himself; that is a dream-like condition, very hard to cure" (C 27:264).

¹⁹ In Collins (reflecting notes originally from 1774–1777), where Kant never explicitly formulates the categorical imperative, he stresses the importance of believing that the rational awareness of what is intrinsically right and dutiful can move us to act. At times, however, he seems to doubt whether such rational awareness alone is sufficient for motivation. On this point, see Schneewind 1997a, xvi. Although Kant may waver on this point in Collins, his later discussions in *Mrongovius II* (1784–1785) and *Vigilantius* (1793–1794) are decisive that virtue involves acting dutifully from duty.

the sake of duty, is *indoles erecta*. That act alone is morally good, which is done because it is a duty. (M II 29:624)

The fact that Kant treats virtue and duty as moral concepts that cannot be prised apart underlies his claim that there is only one (purely) formal principle of ethics or general principle of virtue (one obligation of virtue), namely, the duty to do all of our duties from duty (V 27:541).

An important consequence of this account of motivation in the virtuous person is that moral obligation involves the necessitation of the will by practical reason. To act from duty is to act from a voluntary feeling of respect for the moral law, where acting from such a rationally grounded feeling necessarily involves self-constraint and moral compulsion, and signifies that we obey the law in the face of opposition from inclinations (M II 29:617; V 27:518–520). Kant reasons that, *if* morality ceased to involve the necessitation of the will by practical reason, we would no longer experience moral laws as imperatives, or have any need for virtue, as a steadfast disposition to fulfill one's duties strictly:

Where there is no necessitation, there also no moral imperative, no obligation, duty, virtue, ought or constraint is conceivable. Hence the moral laws are called *laws of duty*, because they presuppose an agent subject to impulses of nature. (V 27:489)

It is this key aspect of Kant's account of virtue – that abiding by moral laws from a purely moral disposition involves doing one's duty from duty and a feeling of necessitation by practical reason – that accounts for Kant's fundamental disagreement with Schiller. Readers familiar with Schiller's critique of Kant's rationalist moral theory in "On Grace and Dignity" will recall that Schiller claims to agree with the fundamental tenets of Kant's moral theory. His concern is primarily with the manner in which Kant presents his doctrines. According to Schiller, the problem with Kant's presentation of his moral theory is that it leaves us with the harsh impression that inclination is "a very suspicious companion, and pleasure a dangerous auxiliary for moral determinations" (GD 169; 205).²⁰ Schiller, however, thinks that this rigid Kantian picture can be improved upon by attempting to "uphold, in the realm of appearance and in the actual practice of moral duty, the claims of sensibility, which are completely repudiated in the realm of pure reason and in moral legislation" (GD 169f; 205). In terms

²⁰ "On Grace and Dignity" was first published in 1793 in the second edition of *Neue Thalia*. References to this essay (GD) are to Schiller 1966 and to the corresponding page in the anonymous English translation of 1902.

of Schiller's governing metaphor, this emphasis on the aesthetic dimension of morality will show how Kantian dignity (*Würde*) is complemented and perfected by grace (*Anmut*).

For our purposes here, the crucial aspect of Schiller's critique to keep in mind is that he thinks virtue requires cultivating one's sensible nature to harmonize positively with reason so that the virtuous person is inclined to act in accordance with the moral law and takes pleasure in moral action.²¹ Schiller conceives of the person who has unified her sensible and rational natures as a "beautiful soul," where a beautiful soul is considered a "child of the house" of practical reason, as opposed to its "servant" (GD 172; 208). In light of the fact that a child of the house has trained her inclinations to accord with duty and to work toward bringing about morally good ends, she is not required to consult reason for guidance before she acts and can even "abandon herself with a certain security to instinct" (GD 173; 209). Schiller argues, further, that this beautiful soul does not experience the moral law as an obligation signifying compulsion and constraint (GD 172; 208). In short, Schiller denies that moral laws necessarily take the form of categorical imperatives for all finite rational beings on the grounds that the imperatival form of the law would strike a child of the house as "a foreign and positive law," a law imposed by an external authority (GD 172; 208). He holds that the person who displays grace is not the one who forces herself to act dutifully, relying on the thought of duty as an incentive, but the one whose inclinations have been so completely transformed as to elicit the right conduct without being commanded, and without experiencing the moral law as an imperative at all.

Kant's published reply to Schiller appears in a famous footnote in *Religion* (RGV 6:23f n.). Yet, while Kant's reaction to Schiller in *Religion* is ambiguous, an important discussion in the Vigilantius' lecture notes on ethics is especially instructive in explaining why Kant rejects Schiller's account of the beautiful soul as an exemplar of moral virtue.²²

²¹ Schiller identifies virtue as "nothing other than an inclination for duty," and he insists that, "unless obedience to reason is a source of pleasure, it cannot become an object of inclination" (GD 170, 169; 206, 204).

²² In *Religion*, Kant states that he and Schiller agree on the most important principles (RGV 6:23n.). Kant then explains that he explicitly distinguishes dignity from grace in an effort to emphasize the purity of duty, which involves "unconditional necessitation, to which gracefulness stands in direct contradiction" (RGV 6:23n.). Nevertheless, Kant says that Schiller's graces are relevant when we turn from an analysis of the nature and ground of moral obligation to consider the aesthetic character of virtue, that is, the feelings associated with virtue as a character trait and the feelings that virtue elicits in us as rational appraisers. In acknowledging that the graces accompany virtue, Kant thus emphasizes the idea that the temperament of the virtuous person is joyous and that the virtuous person does her duty with a cheerful heart. Yet, as commentators have noted, Kant overstates the matter when he

In *Vigilantius*, Kant prefaces his discussion of Schiller by explaining that, for a holy will, what is objectively necessary is also subjectively necessary, which means that a holy will does what the moral law requires without having to be commanded. By contrast, for a finite rational will, whatever action the laws of reason command “does not always result from them, owing to its subjective contingency, rooted in the impulses of human nature” (V 27:489). As we have seen, virtue presupposes some element of struggle and (inner) conflict because our very condition is one in which we must overcome a deep-rooted tendency of will to act on sensible feelings, desires, and interests that run contrary to the moral law. This means that we can never get beyond the bare possibility of being tempted to act contrary to duty and hence we require virtue, as an acquired moral strength of will to constrain or compel ourselves (in accordance with duty) to do what we would not do of our own accord (by our very nature). Whereas Schiller thinks that the moral law takes the form of an imperative only insofar as an agent *actually has contrary inclinations that pose temptations*, for Kant, the imperatival form of the moral laws presupposes only *the possibility for transgressing the moral law, not actual temptation*.

Furthermore, given that moral imperatives entail a notion of necessitation of the finite will by practical reason, Kant remarks that it is “contrary to the nature of duty to *enjoy* having duties incumbent upon one” (V 27:489). On the contrary, we obey the law reluctantly, that is, in the face of opposition from a formidable opponent, an opponent that virtue as “a steadfast determination in obeying the moral laws” enables us to overcome (V 27:490). As Kant thus makes clear in his discussion of Schiller, we require a power of self-constraint and moral compulsion in our efforts to abide by duty from duty, and this fact about our moral experience rules out Schiller’s idea that the fulfillment of moral laws “also has a certain charm (grace) about it” (V 27:490).

Yet, while Kant reminds us that the moral law issues its commands without attracting us, he denies that law demands our “respect in the manner of painful or despotic commands” (V 27:490). The feeling of respect the moral law elicits in us is akin to a feeling of awe and sublimity

suggests that this puts him in full agreement with Schiller on matters of principle. Schiller’s contention is not that grace *accompanies* virtue, but that grace is at least *partially constitutive* of virtue. It is therefore not enough for Kant to agree that the virtuous person has a “heart joyous in the compliance of duty,” because Schiller’s point is that pleasure engages moral action in the genuinely virtuous person and that virtuous person is inclined to do her duty, which she does not experience as an obligation (RGV 6:23n.). Thus, in spite of Kant’s conciliatory tone in his official reply to Schiller, there are substantive points of disagreement between them, which *Vigilantius* helps clarify.

in response to the unconditional dignity tied up with our own self-legislating rational nature. For Kant, then, pleasure factors into our experience of virtue, but the uniquely moral pleasure associated with virtue is always consequential upon doing one's duty from duty (V 27:490).

In sum, Kant rejects Schiller's conception of the beautiful soul as an appropriate model of moral perfection for finite rational beings because what defines our status as imperfect beings is that we have sensible needs and inclinations that do not of themselves accord with the moral law. More precisely, we have a tendency to grant priority to sensible needs and inclinations that run contrary to duty because of our propensity for evil.²³ As a result, the best that we can attain is virtue, a morally good disposition to do one's duty from duty and to compel oneself to act in accordance with the moral law from respect, out of necessitation by practical reason.

Conclusion

As this analysis of Kant's account of virtue in the *Lectures on Ethics* has shown, Kant thinks of virtue as a form of moral self-mastery or self-command that represents a model of self-governance he compares to an autocracy. In light of the fact that the very concept of virtue presupposes struggle and conflict, Kant insists that virtue is distinct from holiness and that any ideal of moral perfection that overlooks the fact that morality is always difficult for us fails to provide an appropriate model of human virtue. No matter how morally good we are or become, virtue remains a disposition to do one's duty from duty, out of necessitation by practical reason. Yet, even though finite rational beings require a power of self-constraint in accordance with the commands of duty to comply with the law (which we obey reluctantly), the virtuous agent displays a unified soul that is at peace. This picture of virtue uncovered from the *Lectures on Ethics* thus reveals the way in which Kant's conception of virtue accords with his foundational commitments in his moral theory, while at the same time representing a more complex theory of moral character and a life lived in accordance with practical reason.

²³ Schiller realizes that Kant's view that we have "radical tendency" (*radikalen Hang*) to act in opposition to the moral law supports the idea that moral laws always take the imperatival form for even the person with full virtue, but Schiller rejects this doctrine of evil and is consequently not worried about the risk involved in allowing moralized inclination to execute morally good conduct (GD 172; 208).

CHAPTER 14

Love

Jeanine Grenberg

Introduction

In reviewing his lectures from the earliest (1760s) recorded by Herder through his latest (1790s) recorded by Vigilantius, Kant's thoughts about love are always placed in a broader context than simply the positive attitude one brings to another person. Kant considers a broad range of kinds of love and objects of love, and considers how these loves relate to a variety of other attitudes, esteem and duty prominent among them.

I do not intend, in this chapter, to provide a complete overview of Kant's reflections on love. Rather, I seek to articulate Kant's suspicions about love of others as a moral motive, and then assess the extent to which, despite these suspicions, he welcomes love into his ethic of duty. We shall discover that Kant's suspicions about love are grounded in worries about how love of self can corrupt love of others, making it illusory and useless. Although Kant will briefly flirt with a proper self-love or well-being that would ground proper love of others, ultimately, in the movement from the Herder to the Collins lectures, there is a solidification of a principle-based ethic of duty grounded in mastery of and esteem for (instead of love of) the self, one which carefully constrains the extent to which love for others is integrated as a moral ideal. Despite these constraints, these lectures, when compared with his published works, represent a high-water mark for the extent to which Kant welcomes love – especially love as a feeling – into his ethics.

1 Herder

We can begin by investigating Kant's earliest Herder reflections, a point from which we will be able to trace Kant's development on both the nature of love of self and others, and the ethic which grows up alongside these thoughts on love.

In the Herder lectures, love of others is presented as unstable and easily corrupted, a state which is in tension with one's efforts to utilize one's own

will well. The first challenge to willing well while loving emerges when Kant suggests that loving others naturally involves “an inclination to please” them. Although this inclination to please another “bring[s] men together,” it is also “a slippery quality,” since it easily turns into “complaisance,” a tendency to “adjust myself to the will of another,” thus “show[ing] that a man has no will of his own, and is without moral worth” (H 27:53). In seeking to love others, we seek the approval of others; we want them also to be pleased with us. But in the effort to accomplish our goal of pleasing others, we ignore ourselves. We fail, in other words, to be true to our own capacity for willing, or to have a “will of [one’s] own.”

This tendency to collapse one’s will in an effort to be pleasing to others is not, however, Kant’s only, or even his largest, worry about love of others. To reveal the more paradigmatic, and problematic, tensions between love of others and the pursuit of “self-will” (H 27:53), Kant presents a contrast between humans in their current, civilized state as opposed to their original natural state.

Kant first envisions “a man of nature” (H 27:64) who finds it easy to love because neither his loving nor his instincts more generally (we apparently cannot yet say that he has a “will” as such) are infected by the demands of society. His “love is *real*” precisely because he “loves *in a more limited way*” (H 27:65). The reality of this love is, furthermore, revealed in its “practical[ity]”; that is, it is a love that actually issues in some good for others, once again precisely because it has limits already built into it by the circumstances of the un-civilized world:

[T]his truly practical instinct is directed, not so much to the increase of good, as to the *prevention of great and sudden harms*; and as soon as they are too much for his powers, wishings and pityings strike him as too foolish; he would have to divert his attention *from himself*, and so he is perfectly ready to turn his thoughts elsewhere. (H 27:65)

In an uncivilized state, one’s love for others is necessarily limited to helping others when “great and sudden harms” present themselves as threats to another’s survival. This turns out to assure a good limit on love because our concern for others is naturally focused where it can achieve the most practical good. Furthermore, such limited love allows one simultaneously to devote adequate attention to oneself without having one’s attention toward others “divert . . . attention” from oneself too much. The man of nature is able to hold onto his own instincts (what we might call his proto-will) even as he expresses a limited, practical love for others.

Even when Kant moves from the “man of nature” to more current moral concerns, he suggests that a *limited* love for others is the best and, ultimately,

most morally praise-worthy state. The need to be true to one's own will, that thing most important for being a moral person, demands that we *separate* ourselves from others and *limit* our connection to them. "In trifles," Kant says, "self-will tends to separate us . . .; but in morality it is worthy of praise" (H 27:53). I must claim a will of my own (a "self-will") in order to become a moral person. But when I claim my will, by that very act I distinguish myself from others, perhaps even deciding that I do not want to please them (or, at least, that as long as I am following my own will properly, I can be indifferent to whether my acts are pleasing to other persons). Kant applauds such firm, willed "indifference, as a moral quality" (H 27:54).

One reason for approving of indifference is that it restrains sympathetic feelings that might encourage excessiveness in our responses to other people. Although indifference is "the opposite of human love," and might even be called "cold-bloodedness," this indifference is also "a very good trait, if it holds the love inspired by sympathy in check, and gives it due measure" (H 27:53). Cold-blooded indifference assures instead that love is "measure[d]." Kant thus encourages us, both from the perspective of the "man of nature" and of current moral concerns, to limit love of others, and find moral worthiness in indifference. When love of others avoids the excess of sympathetic feelings, and is instead properly limited by self-willing, the proper balance of concern for the moral self and concern for others first revealed in the man of nature is assured.

But the movement from nature to civilization presents challenges in implementing this limited love for others. "In present-day civil society," where potential harms are not as present, and where "needs multiply" as society starts to tell us that we need more and more (H 27:65), love for others loses the proper limits it had in the state of nature, and becomes excessive, corrupted and illusory. Here, though, we find a new challenge for one's willing. Earlier, Kant said we give up our own wills in an effort to please others. But now, rather than abandoning my willing, in a wishful effort to maintain it even as I seek the love of others, I instead ground love of others in deceptive, frivolous and ultimately useless willing:

Excessiveness in the way of life . . . creates such wishes and yearnings, and is not good, since it is (1) *useless*; (2) *deceptive*, in that it squanders time and actually impedes practical love; for the love that is too little practical has the love that is all-too-greatly fanciful as its cause. (H 27:64)

This is a particularly interesting passage. Kant suggests that love for others unconstrained by the demands for survival becomes "deceptive" and "fanciful." This, in turn, makes love "useless," nonpractical, the very opposite of

the love of the man of nature. Instead of being focused on real needs of others, love in the civilized state focuses instead upon satisfying “a yearning of the fancy,” and becomes “a topic of romances such as Fielding’s, etc., that has no effect” (H 27:65). Our love for others is no longer honestly focused on or limited by the survival needs of others, and thus loses its practical focus. Not the needs of the other, but instead one’s *own* less compelling needs – these “wishes and yearnings” – infect the would-be love relationship, and prevent love from finding its proper object. An “all-too-greatly-fanciful” love causes, ultimately a “love that is too little practical.”

We should, however, dwell a bit more on what exactly it is that causes this degeneration of love. In the earlier passage, Kant suggests that “excessiveness in the way of life” introduced by civilized society creates illusory, wishful loves. But looking at the Herder lecture notes more closely, we see that these societal encouragements take hold in humans only because something made them receptive to just this corruption. Recall that the human’s first strong desire is to be “pleasing to others” (H 27:53). This is, at bottom, not so much a desire to love others as to be loved by others. But this hope of love from others is, in the end, only one expression of an even more self-serving motivation. The human’s deepest natural concern is to get what she needs to preserve herself, and even to assure that she *gets* more than she *gives* in her interactions with others:

Nature has not framed us to be generous, but to be self-sustaining; sympathetic, indeed, to the woes of others, yet in such a way that the sum shall not be zero; that I not sacrifice as much as I redeem, but preserve myself and my kind. (H 27:66)

Here, then, is the problem for love that emerges: whether naturally, or in society, humans are the kind of beings who want, most centrally, to get what they need for themselves. Yet we also want to be pleasing to others, even hoping that others will act to help us either to preserve ourselves when threats to our survival emerge, or to get the fanciful things society encourages us to desire. Trying to identify a true and authentic love *for another* within this context becomes difficult. On this picture of humans, the connection I make to others is based on hopes *for myself*, not others. It is *my* hope for survival, *my* hope to get what *I* need, that leads me to seek to love others and, more importantly, to be loved *by* others.

We can thus approach Kant’s assertion of the degeneration of practical love into deceptive and illusory love from a new perspective: what is illusory about love in society is that what is presented as love for another is in fact, at its root, most basically a love of self, that is, a desire for self-preservation.

The “wishes and yearnings” created by the excessiveness of society are wishes and yearnings that attach to the self’s concern *for* itself, hopes for what pleasures, securities, or benefits one might gain from others. But the interest in self-preservation and the hope for attainment of things for the self leads humans to present these wishes and yearnings *as if* they expressed love *for another*! That is what is illusory, and ultimately, nonpractical, about love. The state of the purportedly loving person is not at all a state of genuine concern for another person or her genuine needs; it is, instead, a concern for how that person can satisfy a need of one’s own. The true obstacle to making sense of love for others is the excessive, society-encouraged love for self that interferes with our relationship to others. Excessive, illusory and fanciful loves, all ultimately veiled expressions of love of self, thereby stand as *the* obstacle to a proper, authentic love of others. As Kant concludes, “by excess we multiply in fancy our own needs, and thus make practical love difficult, i.e., *eo ipso rare*” (H 27:64).

The first lesson on love from the Herder lectures is that the willing, self-loving self is an obstacle to love of others. The genuine need for the self to will as an individual, and the illicit and excessive tendency of the self to construct and will fanciful and deceptive objects of love to satisfy its own longings, combine to prevent the would-be-lover from successfully loving others as others.

It thus should come as no surprise that, when Kant wants to make sense of a genuine love for others in these lectures, he insists that the first step to proper love of others is a proper love of self. Such an assertion is, by the way, further evidence that the real problem for loving others is grounded not simply in society, but also in the entrenched attitudes one takes toward oneself. If society were solely to blame for our failure in loving others authentically, Kant would have encouraged us either to leave or reform society as an antidote to our problems. But instead, he encourages us to reform ourselves:

[L]et us note (1) that a person does not actively love another until he is himself in a state of well-being; since he is the *principium* of the other’s good, let him first better himself. He should be at ease with himself, and thus the more there is of excess, the less there is of practical human love. For by excess we multiply in fancy our own needs, and thus make practical love difficult, i.e., *eo ipso rare*. (H 27:64)

Kant’s point here is that love of another person requires the would-be lover already to be in a stable state of “well-being.” Without that, she cannot love another; for, as he suggests elsewhere, “self-love must take precedence, since

the love for others simply rests upon it" (H 27:53). Furthermore, Kant suggests that well-being is undermined by the excesses introduced by societal urgings. The way to achieve both well-being and a proper love of others is thus to reject the excesses that society would encourage, those excesses that made our loving of others merely deceptive and fanciful exploits of self-love. As Kant notes here, the real problem with such deceptive loving is that it destroys the proper "practical" nature of true love. For love to be love, it has to have the effect of being helpful to the other person. But a love that is secretly focused only on the multiplying needs of oneself cannot possibly see the other clearly so as to identify and respond to an actual need of the other. Thus, as Kant puts the point, "the more there is of excess, the less there is of practical human love." The rejection of such fanciful and deceptive excesses and the establishment of well-being in oneself is thus the antidote to the discontents of civilization that Kant has introduced: the more that we can overcome the excesses and deceptions that make us present our own needs surreptitiously as a love for others, the more that we can hope to be "at ease with" ourselves and thus begin, from a stable basis, to engage in true practical love *for* others. When well-being is achieved, the human tendency toward excessive, deceptive fancies will dissipate, and a genuine "practical human love" for others can emerge.

There is something attractive about the thought that a good lover is one who is happy and at peace with herself. Nonetheless, one might question whether such a state of "well-being" – that is, having most or much of what one needs, and therefore being at ease with oneself – is an adequate antidote to the tendency toward self-deception about fanciful loves that was the problem for loving others in the first place. It is, after all, the case that excessive wishings and longings developed in humans initially precisely when one's own most basic survival needs were already met. Such satisfaction of need did nothing for preventing fanciful excesses then, so one wonders why Kant thinks that acquiring a state of well-being would accomplish that now. Kant himself provides no further hint of how we should construe well-being such that it would act as an antidote to excesses of self-love.

We will consider these questions further as we continue to study Kant's lectures. For the present, though, we see from this brief review of the Herder lectures that excesses encouraged by civilization, but rooted in concern for self-preservation and self-love, corrupt would-be love of others into a deceptive and fanciful pursuit of the satisfaction of the needs of the self. If we are to reclaim the proper, limited, cold-blooded, practical love of others first displayed by the precivilized man of nature, we must first assure a well-being in oneself.

2 Collins I

We want, ultimately, to articulate more precisely what this managing of one's attitude toward self is that precedes and grounds proper love of others. It would behoove us, however, to preface that pursuit with a previous one: reflection upon the later Collins lectures show that Herder lectures' concerns about both the corruption of love of others by excessive love of self and the ideal of a limited, even cold-blooded, practical love of others remain present in Collins. As we shall see, both these concerns are not only present, but intensified, in the Collins lectures.

Of particular interest in the Collins lectures are two new subsections (in the [first section](#) of the lecture notes titled "I. OF DUTIES TO ONESELF") devoted to questions about love of self ("Of Self-Love", C 27:357ff) and proper mastery of self ("Of Self-Mastery," C 27:360ff). These sections preface both an extended consideration of duties to oneself and then, later, a large section dedicated to duties toward others ("II. OF DUTIES TOWARDS OTHER PEOPLE", C 27:413ff). That this extended discussion of self-love, self-mastery and duties to self comes before any discussion of love of others and duties to others suggests that the original ordering of such things in the Herder lectures (i.e., that concerns about self must be resolved before approaching love of others) remains here also in the Collins lectures.

Details of these lectures' notes confirm this initial impression. In "Of Self-Love," Kant suggests that improper love of self – now distinguished as either *philautia* ("moral self-love") or *arrogantia* ("moral self-conceit") (C 27:357) – corrupts one's moral state overall, both in relation to love of self and love of others. Furthermore, as in Herder, tendencies toward both inactive and deceptive love for others are the fruit of these improper self-loves. Let us consider each improper state of self-love in turn.

First, *philautia* is the less offensive of a still improper state of self-love. It is "only an inclination to be content with one's perfections" (C 27:357). It consists in "mere wishes, and is otherwise inactive" (C 27:359). *Philautia* is thus an unthinking admiration for oneself: it is not so much that I lie to myself about whether I am worthy of love; it just never occurs to me to judge myself in the first place. Instead, I happily, and unthinkingly, lose myself in an unreflective, wishful "self-satisfaction" (C 27:357). This state is nonetheless "contemptible" because it takes its empty, inactive wishes to be good things, but is, at bottom "that whereby the heart only becomes flabby," and "inactive" (C 27:359).

We see in *philautia* an echo of the Herder lecture concern that improper, fanciful and merely wishful self-love undermines active, practical love for

others. Through failure of reflection on my own state, I fail to focus my concern for morality, and for others, properly. I'm already so pleased with what a good person I take myself to be, I never even think about what I have failed to do for others. As Kant puts the point:

[*Philautia*] arises when a man holds his dispositions to be good ones, and thinks by empty wishes and romantic ideas to promote the welfare of the world; he loves the Tartar, and would like to practice kindness toward him, but gives no thought to his closest neighbours. (C 27:359)

Thus far, self-love, as described in Collins, closely parallels what we have seen in Herder. *Arrogantia*, however, goes a step further in recognizing the moral repugnance of a certain kind of self-love. Unlike *philautia*, *arrogantia* is an "active" state (C 27:359), but this does not make it a good thing. It is instead "a far more damaging defect" of self-love than *philautia*, since in it, "the measure of the [moral] law" is misapplied entirely (C 27:358), resulting in "an unwarranted pretension to merit" for the self (C 27:357). Whereas the person of *philautia* unthinkingly loses herself in her unreflective, fanciful beliefs about herself, leading to a flabby, sentimental inactivity, the person of *arrogantia* is indeed active, albeit active toward the wrong, even "evil" (C 27:359) ends, all grounded in a deceptive, offensive and unwarranted – but principled – self-confidence in one's moral superiority.

In a new twist beyond Herder on how exactly self-love can go wrong, both *philautia* and *arrogantia* are, furthermore, a result of weakness in the way we assess ourselves morally. For both, the tendency to seek morality in "examples" or "experience" (instead of recognizing the basis of morality in reason itself) results in a misplaced love of self. *Philautia* results simply when someone weighs herself only against moral examples (instead of the more rigorous law of reason) and tends, thereby unduly to assess oneself as "blameless and without fault" (C 27:357). But *arrogantia* results when one more deeply mischaracterizes the moral law: "the moral law is thought of in a narrow and indulgent fashion," and thus "the moral judge within us is partisan" (C 27:357). In other words, the person of *arrogantia* actively deceives herself into believing that she is more moral than in fact she is, whereas the person of *philautia* is blissfully ignorant of any moral failings simply because she has never reflected deeply enough on herself to make her believe that she is anything but blameless. Kant clearly finds the self-deceptive *arrogantia* more deeply problematic. The person of *arrogantia* "persuades himself that he has acted rightly, according to principles. Nothing is more crafty and repulsive than to fabricate such a law for oneself, whereby one may do evil under the aegis of the true law" (C 27:359). The

fanciful and deceptive tendency, seen in Herder, to believe oneself blameless thus takes a nasty turn here and becomes an entrenched, evil and self-deceptive self-love.

Kant emphasizes, however, an interesting similarity between these two forms of self-love. For both *philautia* and *arrogantia*, “[t]he less strictly the moral law is taken, and the less strictly the inner judge passes judgment upon us, the more arrogant we tend to be” (C 27:357). In other words, there is an inverse relationship between our strict attention to the moral law and the impropriety of our self-love: the more strictly we attend to the moral law, the less likely we will fall into arrogance; the less strictly we attend to the moral law, the more likely we will fall into not just mindless *philautia*, but also a deeply-seated, deceptive and potentially evil arrogant belief in our moral stature.

We find a clue here, then, about how to provide a corrective to that excessive self-love that gets in the way both of morality and of authentic love of others: to build a proper basis for love of others, Kant suggests not that the self should acquire a state of well-being (as in Herder), but instead that one should acquire the right relationship of oneself to the moral law. Indeed, as Kant already prefigures in this discussion of self-love, a person with true moral concern will avoid “love” of self entirely, preferring instead to find a positive attitude toward oneself based in a more moral attitude of “esteem,” not love: “Everything in moral philosophy that increases self-love should be rejected, and only that recommended which makes us worthy of esteem [I]f then we are no object of love, we can look anyone in the eye with confidence, . . . for in that case we have worth” (C 27:358).

This replacement of well-being and love of self with concern for the moral law and a resulting esteem for oneself is a significant change from Kant’s Herder lecture discussions of such things. There, he suggested acquiring a state of “well-being” (H 27:64) in order to provide a solid basis for loving others, a state more focused on assuring one’s self-love and happiness than on assuring one’s moral status. This person of well-being needed, after all, to “be at ease with himself,” and even to be able to “better himself” (H 27:64). But now, Kant abandons the idea that loving others is grounded in one’s own well-being. Instead, the secure foundation upon which a love for others should be built is acquisition of the right relationship to the moral law, an accomplishment that grounds “esteem,” not love, for oneself. Kant even suggests that love of self is a disposable thing when he claims that even if “we are no object of love [apparently either of my own love for self or of others’ love for me], we can look anyone in the eye

with confidence” nonetheless, since I still have a “worth” which makes me deserve esteem, instead of love.¹

To appreciate this transformation of self-love into self-esteem, we need to turn beyond the section “On Self-Love” to the [following section](#), “On Self-Mastery.” Here, we find more details about how to compare oneself to morality, a process through which one strengthens one’s moral motivations and thereby acquires that proper esteem for self which serves as a more secure basis for love of others.

At the very opening of this section, Kant introduces a new way to think of the Herder suggestion that some concern for self must precede proper love for others:

The general *principium* of self-mastery was that esteem for one’s person in regard to the essential ends of human nature, and the self-regarding duties, are conditions under which alone the other duties can be performed. (C 27:360)

What is striking here, especially in comparison to the Herder language of “well-being,” is Kant’s more heavily moralized approach to acquiring the proper attitude toward self. Instead of reflecting upon how the self can become “more at ease” with itself or how it can “better” itself (H 27:64), Kant introduces language of “self-regarding duties.” Instead of speaking merely of possible love for others, Kant introduces the idea of “other duties” (presumably, given the context, duties to others, including a duty of love) which will be accomplished only to the extent that previous “self-regarding” duties are accomplished. All this language of “duty” is, of course, familiar to those who have read Kant’s published works on morality (and this is a good point to remind our readers that the Collins lecture notes come mostly from 1784, a time when the *Groundwork* had already been written but not yet published).² But, compared to Herder, this language of duty, and the replacement of love for self by “esteem for one’s person” and “self-regarding duties” is new, even unexpected.

¹ This rejection of any moral form of self-love will come as a surprise to readers of the *Critique of Practical Reason* because, there, even as Kant warns of the excesses of self-love, he admits a “rational self-love” (KpV 5:73) as warranted. That was not published until 1788, so one might imagine that, after a brief flirtation with abandonment of self-love, Kant returns to this more balanced position. Interestingly though, in the Vigilantius lectures (from 1793–1794) Kant, after briefly suggesting that love of self “only becomes a fault . . . when it excludes others from our love or inclination” (V 27:620), instead reverts to his Collins insistence that the proper attitude toward self is not love but esteem: “[L]ove of well-liking towards oneself is differentiated and transformed into *self-estimation of oneself*. If it rests on a prior close examination of oneself, it not only differs from self-love, namely well-wishing, but is also self-justifying in its own right” (V 27:621).

² See Schneewind [1997a](#) for more details.

Of course, in discussion of Herder, I had already expressed some skepticism about whether achieving a state of well-being was the proper antidote to fanciful excesses of self-deception. Perhaps Kant himself, between the Herder and Collins lectures, had a similar turn of mind. If the thing to be cured to assure an ability honestly to see and respond to the true needs of others is our own tendency to disguise our own needs as love for others, then it seems that just getting into a state of well-being isn't a strong enough cure. I can, after all, have most all my needs met; but, instead of stopping needing things, I could just keep manufacturing more needs. This was, after all, Kant's own worry about tendencies in civilized society. We thus now have a stronger, more effective – and moral – cure: impose upon yourself moral demands that refuse to allow you such fanciful deceptions under any circumstances, and that simultaneously provide a firm basis for accepting yourself that is not grounded at all in the question of whether you get what you want or not.

What is now required in order to attain that proper attitude toward oneself which will promise the possibility of fulfilling duties to others (including, presumably, a duty of love to others) is that one must acquire not love but “mastery” of oneself, a mastery which will entitle oneself to “esteem for one's person.” Such mastery is, furthermore, acquired only when one “discipline[s]” and “command[s]” “a certain rabble element” of one's nature, viz., “sensibility” (C 27:360). That is, one must attend carefully to oneself so as to assure that the temptations of sensibility have not perverted one's actions and disposition, and so ultimately to assure that moral feeling is the strongest aspect of our sensible natures:

[W]e first have to discipline ourselves, i.e., to root out, in regard to ourselves, by repeated actions, the tendency that arises from the sensory motive. He who would discipline himself morally must pay great attention to himself, and often give an account of his actions before the inner judge, since then, by long practice, he will have given strength to the moral motivating grounds, and acquired, by cultivation, a habit of desire or aversion in regard to moral good or evil. By this the moral feeling will be cultivated, and then morality will have strength and motivation; by these motives, sensibility will be weakened and overcome, and in this way self-command will be achieved. (C 27:361)

What we see here is a cure that more obviously and directly responds to the original sickness. In Herder, what got in the way of love for others was a tendency to disguise manufactured needs of the self as love for others. Here, Kant implicitly suggests that such tendencies are based in a common ground, the so-called sensory motive. And now, instead of encouraging

the satisfaction of some of one's sensory motives, in the hope that satisfaction on some grounds will establish a state of well-being and prevent excesses in the pursuit of such ends overall, Kant insists that the real cure is a repeated disciplining of sensory motives. We must "root out" the tendency they ground and replace it with "a habit of desire or aversion in regard to moral good or evil."

Kant's Herder suggestion that one must establish a state of "well-being" before turning to love of others is thus here turned into something else. Instead of building love for others on a solid basis of love for self, Kant shifts to building a solid basis of esteem for self in one's capacity to master oneself morally, especially in one's capacity to master one's unruly sensible inclinations. This is the new, more solid basis for avoiding the deceptive and fanciful self-focused excesses of love of self disguised as love of others.

This concern to replace love of self with an attentive moral esteem for and mastery of self is, furthermore, connected with a Collins reaffirmation and expansion of the moral value of a "cold-blooded" morality initially seen in Herder. In Herder, cold-blooded indifference was seen as a way of properly limiting love toward others by limiting the deceptive, fanciful, and wishful excesses of love. In Collins, we see the same theme; indeed, Kant's appeal to it seems an effort to elevate cold-bloodedness into a new ethic of principle that would replace any merely flabby, inactive and lost-in-good-feeling ethic of love:

A cold-blooded temper towards others is one that evinces no loving affection or stir of emotion. The man to whom such stirrings of good-will are unknown, is cold. Yet cold-bloodedness should not, on that account, be censured. The poets may rejoice at being steeped in warm feeling and affection, and heap abuse on cold-bloodedness; but if it is accompanied by principles and good dispositions, those who possess it are at all events people who can be relied on. A cold-blooded guardian who means me well, an advocate or patriot of that type, are people of steadfast character, who will surely do their utmost on my behalf. Whereas cold-bloodedness, in an evil-doer, is all the worse on that account, in a good man – though it may not sound so well – it is actually better than a warm feeling of affection, since it is more constant. (C 27:420)

Not only, then, does Kant encourage us to abandon self-love in favor of self-mastery and self-esteem. Beyond that, he encourages us to abandon warm-hearted love of others in favor of a cold-blooded and principled – but also active and effective – concern to do well for others. Cold-bloodedness might "evince no loving affection," but it is "actually better than a warm feeling of affection, since it is more constant." A constancy and reliability in actions toward others – what he might have called a more "practical" love in the

Herder lectures – is thus to be preferred to a felt love toward others. We thus see in the movement from Herder to Collins, a hint of how Kant’s now-familiar principle-, esteem-, and respect-based (instead of love-based) ethic emerges from suspicion of the motive of love for others, rooted ultimately in suspicion of the sensible motive of love for self.

Overall, then, in Collins, we see the reaffirmation of two Herder themes: worries about wishful and inactive self-love, as well as the value of a certain indifference or cold-bloodedness in our relations to others. But in Collins, we also see a deepening of both these themes. First, beyond a mere wishful, inactive self-love, Kant worries about a principled but undue love of self that makes for more deep-seated moral problems in our relations with both self and others. Concerns about self-love are thus, if anything, made more explicit and the concerns grounded in them more pressing than they were in the Herder lectures. Similarly, Kant’s brief assertions in Herder about the value of indifference in moral action (as a limit on the excesses of one’s wishings and yearnings) become solidified in Collins into a cold-blooded, principle-based ethic instead of a warm-hearted love-based ethic.

3 Collins II

Given all these difficulties that love of self makes for love of others, it should not surprise us that Kant encounters some difficulty in the Collins lectures when he is faced with the Gospel command that “Thou shalt *love* thy neighbor” (C 27:417), and not just esteem her or be constant toward her! Given Kant’s earlier worries about how both love of self and love of others go so wrong, can he really provide in these lectures any account of morality that truly integrates this Gospel demand to *love* others? Kant has already, throughout both Herder and Collins, expressed a deep suspicion of the motive of love, both as directed to self and others; so, when God demands of us that we love our neighbor, Kant must begin to wade in some deep, tricky waters! In the end, if Kant is going to be true to the Gospel injunction to love one’s neighbor, he must integrate an ethic of love within his newly forming cold-blooded ethic of principle. Let us investigate the extent to which Kant accomplishes this in the Collins lectures.

We have already seen the first step in this process: improper love of others is grounded in improper and excessive love of self. So the first step toward proper love of others is to get one’s attitude toward oneself in the right order. And we saw, in the [previous section](#), that the corrective to the excesses of self-love involves replacing love for self with a moral esteem for

self, itself grounded in attentive and disciplined mastery of oneself, especially of one's sensible motives. Getting this self-regarding attitude right was a "condition . . . under which alone the other duties can be performed" (C 27:360). We have, further, suggested that this mastery of sensible motives gets at the heart of the original problem with self noted in Herder: if we entirely root out our tendency toward excess in self-focus, we no longer have the ground for deceptive ways of pretending to love others, all really just focused on loving the self.

What, then, does a more honest and reliable love of others based in this new self-mastery and self-esteem look like? Interestingly, although we find a new basis for love of others over against the Herder recommendation of self-love, the recommendations for proper love of others that result from this firmer basis of self-esteem actually look very similar to those found in Herder. In Collins, Kant recommends that, at the heart of commanded love of others is what he described in Herder as "practical love." Consider, for example, the following passage, in which Kant speaks of what he now calls "practical desires":

[W]e must see to it that our inclination to love the other, and wish for his happiness, are not idle longings, or desires with no outcome, but practical desires. A practical desire is one that is directed not so much to the object as to the actions whereby this object is brought about. We should not only take satisfaction in the welfare and happiness of others, but this satisfaction should relate to the effectual actions that contribute to this welfare. In just the same way, I should not wish, when the other is in misery, that he might be rescued from it, but should attempt, rather to rescue him . . . [S]o far as I can extend a hand to help him, I am to that extent able to promote his happiness, and sympathize with his plight; but I show no sympathy whatever for his plight in harbouring passionate wishes for his deliverance. The heart, then, is only a good heart insofar as it is able to contribute something to the other's happiness, and not when it merely wishes for that . . . [T]he only one to have a kind heart is he who contributes something to that happiness. (C 27:421)

This passage is reminiscent of Kant's Herder lecture worries about love of others: love for others is most easily undermined by a self-serving fascination with "passionate wishes" that accomplish nothing. One wants to feel good instead of really doing the hard work of helping another person. Although he does not state it explicitly here, it is reasonable to think, given his earlier discussion of self-mastery, that Kant now finds such excesses best controlled by mastering one's sensible motives and esteeming the self properly. When

we set such excesses aside, we are more capable of “contribut[ing] something to the other’s happiness” instead of just feeling good about ourselves.³

Kant’s continued commitment to a certain “cold-bloodedness” in the realization of the Gospel demand to love others is also evident here. This love is not so much a felt love as it is an active, practical love: I succeed in “sympathiz[ing]” with the “plight” of another only to the extent that I succeed in “extend[ing] a hand to help him.” Indeed, at other points in the text, Kant explicitly rejects the idea that we could be commanded to feel a certain way toward specific individuals: “Well-wishing from love,” says Kant (a “well-wishing” defined as “the wish and inclination to promote the happiness of others”) “cannot be commanded” (C 27:417). As such, the command to love others as individual others is best accomplished when “love” is understood not as a feeling for others but is focused on practical action which succeeds in doing something good for others.

One might expect Kant, given his coldblooded ethic, to stop at this point. He has defended a sense in which the Gospel command to love your neighbor can be integrated within his principle-bound ethic: instead of commanding a feeling of love for one’s neighbors, the Gospel command is a command to do genuinely good things for one’s neighbors. Yet, interestingly, Kant does continue to struggle with the question of whether some more feeling-based (as opposed to action-based) love of others is also commanded as a duty.

First, Kant suggests that, although “well-wishing from love [i.e., “the wish and inclination to promote the happiness of others”] cannot be commanded,” we may find that when “we do well by someone from duty, we get used to this, so that we subsequently do it from love and inclination as well” (C 27:417). As such, in a precursor of a position that does not arise in his published texts until the *Metaphysics of Morals*, Kant

³ Kant has perhaps not reflected when he made the shift from grounding love of others in love of self to grounding love of others in esteem for self that the Gospel demand to love one’s neighbor is most fully articulated as the demand to love your neighbor *as you love yourself*. Clearly, if self-love of any sort is rejected, then however it is that Kant resolves the Gospel demand to love others, he will not be able to satisfy this aspect of it. If the moral person is not to love herself at all, then love of self provides no model or guide for understanding how we should love others. Interestingly, late in the *Vigilantius* lectures, Kant seems to want to correct this failure to account for the full Gospel command, and provides a rather obscure rendering of how we can understand love of others as grounded in love of self: “Now from the love for God, the second commandment, to love thy neighbor as thyself, is known by inference. God loves all men, i.e., He makes their welfare His end; now in that we make this end of His into our own, we love all men, and since this consists in observing all our duties, and thus primarily the duties to oneself, we love ourselves for the sake of other men, or others as ourself; the self-regarding duties are thus indirectly included here, and the source of all love is the commandments given to us out of God’s goodness, which = our duties” (V 27:722).

suggests that “even love from inclination is a moral virtue, and might be commanded to this extent, that one should first practice well-doing as a duty, and later, through habituation, out of inclination as well” (C 27:417).

But Kant continues to struggle especially with the question of whether both well-wishing love (“the wish and inclination to promote the happiness of others”) and well-liking love (“the pleasure we take in showing approval of another’s perfection”) (C 27:417f) can be commanded more directly as duties. Kant manages the former by suggesting that I am indeed commanded to have a well-wishing love for others, but that this command consists more precisely “not . . . in merely wishing someone well, but in wishing that he might also be worthy of it, and that sort of well-wishing love we may also have for our enemies. Such well-wishing can always be heartfelt” (C 27:417). If someone seems not to deserve my love, I can coax my feelings by envisioning this person as someone striving toward becoming love-worthy. Kant thus, despite his assertions both in *Collins* and his published works that a feeling cannot be commanded, commands this sort of feeling of love.

There is, however, a deeper problem with the command of well-liking love:

[W]ell-wishing love to one’s neighbor can [thus] be enjoined upon everyone. But well-liking love to one’s neighbor cannot be generally commanded, since nobody can have such a liking where there is no object of which to approve. (C 27:418)

The problem here is that we could hardly be expected to take pleasure in another’s perfections (the definition of “well-liking love”) when that other *has* no perfections! We get around the problem of the command to have a well-wishing love (i.e., to have “a wish and inclination to promote the happiness”) of imperfect others by changing the inclination into one pointed toward their eventual worthiness. But although we can wish for the eventual worthiness of another, we cannot actually take pleasure in only the eventual worthiness of another. So, Kant cannot handle a command for well-liking love in the same way.

Kant now makes an interesting move: although I need to *act* lovingly for specific individual others (perhaps even, or especially, those imperfect others for whom it is difficult to feel), and I am even commanded to have a well-wishing *feeling* of love for specific individual others (especially for their eventual moral worthiness), it is *not* commanded of me that I feel well-liking love for specific individual others. Instead, feelings of well-liking must be directed at the “humanity” we find in others:

There is, however, a distinction to be drawn in a man between the man himself and his humanity. I may thus have a liking for the humanity, though none for the man. I can even have such liking for the villain, if I separate the villain and his humanity from one another; for even in the worst of villains there is still a kernel of good-will . . . If I now enter into his heart, I can still find a feeling for virtue in him, and so humanity must be loved, even in him. Hence it can rightly be said that we ought to love our neighbours. I am not only obligated to well-doing, but also to loving others with well-wishing, and well-liking, too. (C 27:418)

Kant thus, perhaps surprisingly for those familiar with his limited integration of commands for feelings of love discussed in the *Groundwork*, the *Critique of Practical Reason*, and the *Doctrine of Virtue*, suggests here that we can be commanded to love others in four distinct senses: (1) to love in the sense of doing something good for another, (2) to love in the sense of cultivating a disposition or virtue of loving others through inclination, (3) to love in the sense of having a feeling of well-wishing toward the actual or potential moral worthiness of another, and (4) to love in the sense of having a feeling of well-liking toward the humanity in all others.

To what extent, then, can we say that love is integrated into Kant's coldblooded, principle-based ethic? In a position more strongly held here than in his published works, love of self is rejected entirely. But love of others, while placed within strict limits, is defended more fully than in his published works. Love is most centrally an action-love, not a feeling-love; but feeling-love *is* commanded, not only as a duty to cultivate a feeling of love over time, but also as precise demands to feel for others, either for individuals as they might eventually be or not for individuals at all, but only for the humanity in an individual. Interestingly, though, this limited integration of love (and especially his integration of feelings of well-wishing love and well-liking love as actually commanded) is the high-water mark for just how much love is integrated by Kant. In his published works, Kant argues that love as a feeling cannot be commanded, even that to command a feeling is "an absurdity" (MS 6:401); but in Collins, he asserts that love as a feeling, constrained in the ways just suggested, can indeed not only be commanded to be cultivated, but can be commanded to be felt.

Love of honor, emulation, and the psychology of the devilish vices

Houston Smit and Mark Timmons¹

Much has been written in recent years about virtue and the particular virtues featured in Kant's ethics, focusing in particular on the *Tugendlehre*, [part II](#) of *The Metaphysics of Morals* (see, for example, Betzler 2008). Less attention has been given to what Kant has to say about vice and the particular vices. Kant's own discussion of select vices in the *Tugendlehre* is quite brief, punctuated with remarks about the psychological sources of vicious character traits. In contrast, what we find in some of the lecture notes is a fairly rich discussion of the psychology of other-regarding vices involving what Kant refers to as emulation – an “impulse” implanted in human nature that strongly inclines humans to be “equal to the other in every respect” (V 27:695). Although this implanted impulse has the good of self-improvement as its purpose, under certain circumstances it contributes to the devilish vices of envy, ingratitude, and *Schadenfreude*. Our primary aim in this chapter is to provide an interpretation of Kant's views concerning the psychology of the devilish vices – how individuals succumb to these vices – based on remarks in the *Lectures*. Of particular interest is the extent to which these vices share an underlying psychological unity, a claim suggested by some of Kant's remarks.

Before turning to the psychology of the devilish vices, we devote the first two sections of the chapter to a discussion of the elements that figure in Kant's conception of virtue, including the dynamics of virtue acquisition and maintenance. Proceeding in this way provides the needed psychological background for understanding Kant's account of the genesis of the devilish vices.

I True love of honor and proper self-esteem

The rudiments of Kant's conception of virtue, important for understanding the dynamics of virtue acquisition, include the related notions of *true*

¹ This is a thoroughly collaborative work; the order of authorship is alphabetical.

love of honor and *proper* self-esteem. Our remarks about these notions in Kant will of necessity be brief; they are not intended to be a full explication of them and their role in Kant's moral theory. Rather, we intend our remarks to be sufficient for purposes of understanding the psychology of the devilish vices.²

In the *Anthropology*, Kant makes a telling remark about love of honor, one that indicates that it is central to his conception of virtue: "Love of honor is the constant companion of virtue" (ApH 7:257). As we understand love of honor it is a virtue that one has a duty to cultivate, indeed Kant claims that it is "the highest duty of humanity to oneself" (V 27:664). It is not entirely clear what Kant means by referring to it as the *highest* of duties to oneself, but on what we take to be a plausible reading, this duty directs one to acquire or, more plausibly, *strive* to acquire a fairly generic character trait that plays an absolutely central role in Kant's conception of virtue, a role, as we shall see, that is not limited to the domain of self-regarding moral concerns. Let us explain.

Kant's generic conception of virtue is (as is well-known) the ideal of a nonholly rational agent who fulfills her moral obligations out of a sense of duty, that is, not only fulfills them, but does so from the sole motivating reason that such and such action or end is required or prohibited by the moral law. In the *Tugendlehre*, Kant defines virtue as "the moral strength of a *human being's* will in fulfilling his *duty*" (TL 6:405). The strength in question concerns the "purity . . . of one's disposition to duty, namely, in the law being by itself alone the incentive, even without admixture of aims derived from sensibility, and in actions being done not only in conformity with duty but also *from duty*" (TL 6:446). This conception of virtue specifies the "formal" end toward which one is to strive: a developed capacity to reliably perform duties from the sole motive of duty.

The most generic "material" or substantive ends that help structure one's efforts to achieve virtue are self-perfection and the (morally permissible) happiness of others. With regard to the end of self-perfection, the virtue Kant refers to as love of honor represents a specification, although still at a fairly high level of abstraction, of the sort of character trait one ought to acquire in striving toward this end. As a trait, love of honor essentially involves a particular attitude toward oneself that Kant refers to as proper self-esteem, the attitude that "underlies true honor" (V 27:667). Proper self-esteem is, one might say, the emotional core of love

² For an illuminating discussion of love of honor in Kant's ethics, including a defense of the claim that it is a virtue in Kant's scheme, see Denis 2014.

of honor as a character trait. So, let us take a closer look at Kant's conception of proper self esteem.

A section in Collins entitled "Of Proper Self-Esteem" begins with the remark that "This self-esteem includes, on the one hand, humility, yet on the other a true noble pride" (C 27:348). Regarding the sort of humility in question, Kant says:

We have reason to harbour a low opinion of our person . . . For if we compare ourselves with the holy moral law, we discover how remote we are from congruity with it. This low opinion of our person arises, therefore, from comparison with the moral law, and there we have reason enough to humble ourselves. (C 27:348f)

As we have noted, the content of the moral law (its requirements) includes a duty of self-perfection – a duty to strive to become a person of virtue, which includes not only fulfilling one's various moral obligations, but doing so on the basis of moral principles that one understands to be supremely authoritative. Again, one is to strive to make the moral law a sufficient motive in fulfilling one's duties. So in this passage the emphasis is on the high ideal of personhood – being a person of overall virtue, a task, which in Kant's view, is "always *in progress*" (TL 6:409). Kant thinks that an honest appraisal of one's conduct and character in relation to this ideal will have a humbling effect on one's self-conception, thwarting not only moral complacency or laziness with respect to the goal of self-perfection but also moral self-conceit, a kind of arrogance. The important point here is that the sole standard for the kind of self-appraisal that Kant refers to as proper self-esteem is the moral law. As we shall see later when we discuss the devilish vices, a failure to understand this true measure of proper self-esteem is central to Kant's explanation of how such vices arise and take hold.

The passage we are examining continues with remarks about true noble pride.

But in comparison with others, we have no reason to entertain a poor opinion of ourselves, for I can just as well possess worth as anyone else. This self-esteem, then, in comparison with others, is noble pride. A low opinion of one's person in regard to others is not humility; it betrays, rather, a petty soul and a servile temperament. . . . [This] self-esteem is reasonable, for we do the other no harm by it, if we deem ourselves equal to him in worth. (C 27:349)

The basis of one's equality with all other human beings is one's dignity as a rational creature, a status that is grounded solely in one's law-giving capacity as an autonomous rational agent. It is something the mere possession of

which is not an accomplishment, but an essential part of one's humanity, which the moral law demands that we respect in ourselves and in others.

So what we find in these two passages is a dual conception of self-esteem. Both conceptions involve the moral law as the sole basis on which proper self-esteem is grounded, and thus the law serves as a standard in relation to which one is to judge proper self-esteem by engaging in an honest moral assessment of one's conduct and character. The sort of self-esteem Kant associates with noble pride is what we might call "recognition self-esteem." All human beings, in virtue of their status as beings with dignity, deserve a certain level of respect: recognition esteem. The dignity in question grounds, both ontologically and epistemologically, proper recognition self-esteem.

The sort of self-esteem concerning the humility that results from comparing one's conduct and character with the ideal of personhood is what we may call "appraisal self-esteem." Complying with the strict demands of duty grounds what Kant refers to as the negative sense of true love of honor. It is negative in the sense that in complying with strict (perfect) duty one avoids being a proper object of contempt. Going beyond the demands of perfect duty and engaging in meritorious actions of beneficence grounds what Kant refers to as the positive sense of true love of honor (see V 27:665).³ Such actions make one a proper object of positive esteem.

Given that a proper sense of self-esteem "underlies," as Kant says, true love of honor, we can note the following. The true lover of honor is someone who recognizes her worth as a being with dignity grounded in her capacity for moral law-giving. This is recognition self-esteem. Associated with this true love of honor is proper self-regard, which results from her comparing her conduct and character against the demands of moral self-perfection set by the moral law, and finding that she more or less measures up to the standard in question. This is appraisal self-esteem. Of course, any such positive self-appraisal must be tempered by the sort of humility that opposes moral self-complacency and moral self-conceit.

Kant contrasts a person who is a true lover of honor with persons who have a mere desire for honor from others, which, in his view, constitutes

³ From these passages, it is not entirely clear whether true love of honor in either the negative or positive senses requires that one's honor-conferring actions be done from the sole motive of duty and thus are not only in accord with duty but enjoy moral worth. Actions that fulfill imperfect duty count as meritorious in Kant's view, regardless of whether they are motivated solely by one's sense of duty. (For defense of this latter claim, see Timmons 2002.) However, as we are about to see (in a passage we quote), in contrast to the true lover of honor, someone whose dutiful actions are motivated merely by the hope of gaining the approval of others is someone whose actions lack moral worth. This suggests that the actions of the true lover of honor are motivated by duty.

a morally objectionable form of ambition. Contrasting the true lover of honor with a “desirer” or mere honor-seeker, Kant says of the former that “He sets out from self-esteem” (V 27:666), while

[t]he desirer of honour, however, does not set out from principles and the steady pursuit of them, but from pragmatic means of making the actions he thereby engages in glitter to the eye, and extort the approval of others, although on scrutiny they lose their moral worth. (V 27:666)

Furthermore, those who merely desire or, as Kant sometimes puts it, “crave” positive recognition from others (C 27:408f), are individuals whose sense of self-esteem is based on a comparison of their moral worth or well-being (or both) with their perceptions of the worth or well-being of select others. This point is crucial for understanding Kant’s explanation of the devilish vices. The true lover of honor is someone for whom the moral law is the sole basis for her moral self-esteem. As Kant remarks in Collins, “moral self-esteem, which is founded on the worth of humanity, must never be based on a comparison with others, but only on comparison with the moral law itself” despite the fact that “People are very much inclined to take others as the measure of their own moral worth” (C 27:349). As we shall see, one’s taking the condition of others as the relevant measure of one’s self-esteem is what leads to love of honor in a “bad sense,” the central character defect characteristic of the devilish vices.

2 The dynamics of virtue

This section will be devoted to the psychological mechanisms, including drives (*Trieben*), instinct (*Instinct*), and impulses (*Antrieben*) that nature has implanted in human beings, which bear directly on coming to develop the true love of honor that is at the center of Kant’s conception of virtue. The three that play a role in Kant’s psychology of virtue (and vice) are: the drive to be respected and the related impulses associated with rivalry and emulation. After briefly introducing these natural endowments, we explain their dynamic in the development of a virtuous character, which includes the need for individuals to have a conception of proper self-contentment.

The drive to be respected

In a section of Collins entitled, “Of the Two Drives of Nature, and the Duties Relating to Them” (C 27:407–12), the drives in question are: *to be respected by others* and *to be loved by others*. Kant claims that the former drive is

typically stronger, and offers two reasons in support of his claim. First, respect or esteem is properly directed toward one's inner worth, while love only has to do with "the relative worth of others" (C 27:407), which presumably means that whether one is lovable depends on how others respond to one's overall personality. Lovability, one might say, is in the eye of the beholder. Second, respect "confers greater security" (C 27:407) in the estimation of oneself by others, as being a proper object of respect is something that one is owed by others, while being an object of love (even a proper object of love) depends, as just noted, on the variable sentiments of others which determine whether they happen to find one loveable. Moreover, being subject to contempt for lack of being respect-worthy is more painful (psychologically) than not being loved or being an object of hate. The significance of the drive to be respected by others concerns its purpose:

Providence has implanted that drive in us, so that our actions and practices might conform to the general judgement of others. For if we lacked it, we would not make our actions so acceptable to the community. We might go astray in our own judgement, so that our opinions would often be much mistaken, if they were to rely solely on our own judgement. Hence this drive leads us to compare our judgements concerning our knowledge with the opinion of others. (C 27:411)

The passage goes on to stress the importance of comparing one's opinions to "universal reason" which is the "tribunal before which our knowledge has to stand" (C 27:411). So, at least part of the import of this drive as it bears on the development of virtue in oneself is epistemological, and is associated with a duty to seek out the judgment of (at least certain) others.

Two questions about this drive and its significance are worthy of address. First, one might wonder about the significance of the drive for respect *from others*, since in other places Kant makes clear that "A lover of honour finds in himself no need to be known (no [craving for celebrity]); he does not require to be highly esteemed by others" (V 27:665). But as Kant makes clear in the Collins passages under examination, the significance of having others recognize one's love of honor grounded in proper self-esteem is for the positive effects it has on humanity generally. In contributing to what Kant calls a "moral bond," everyone should avail themselves of the opportunity to display not only a negative example to others, by doing "nothing evil, but also provide a positive one, in exhibiting an element of good. So our actions must not only be good, but also be seen as an example in the eyes of others. They must spring from a love of honour" (C 27:411f), and in doing so one avoids a mere

appearance of moral goodness, which is often motivated by a mere desire to be viewed by others as worthy of exceptional respect or esteem, and thus a person of moral superiority. Such behavior, including the motives behind it, likely constitutes a kind of unsavory ambition or perhaps arrogance. So the significance of the drive for respect is not merely epistemological; the remarks we have just quoted stress the social import of this drive.

Second, because (as noted earlier) true self-esteem is properly grounded only in a comparison of one's behavior, attitudes, and motives with the moral law and not in comparison of oneself with others, one might wonder whether the epistemological import of the drive to be respected by others is at odds with proper self-esteem. But Kant addresses this question by claiming that when it comes to determining the principle of true honor-worthiness, one is not to rely on others' judgments:

In regard to rectitude, which I perceive by my own reason, I can follow no opinion, but must be guided by my own principle, discerned through reason. But if it is a matter of custom, for example, I have to be guided by the opinion of others. (C 27:412)⁴

Presumably, Kant's point is that the opinions of others have epistemic weight in moral self-assessment only when it comes to particulars having to do with the application of this a priori cognized principle to determine whether, in certain contexts, one's actions properly constitute respect for humanity. Considerations of custom matter because they often determine what does and does not count as, say, an insult or an invasion of privacy. Such empirical issues are important in applying, or being guided by, the basic principles of morality in particular circumstances. In the *Tugendlehre*, Kant devotes a very brief chapter (TL 6:468f) to this issue of application in which he raises questions about how one is to respond properly to individuals who occupy various roles, who have a particular moral character, who find themselves in such circumstances as prosperity or poverty, and so on.

Since becoming (or striving to become) a lover of honor and thus a person of virtue requires knowing one's moral obligations (as well as becoming properly motivated), the epistemological purpose of the drive to gain the respect of others is of some significance. Of greater significance for our topic are the closely related instinct of rivalry and the impulse of emulation.

⁴ Kant says, "In regard to his natural obligations, nobody can be in error; for the natural moral laws cannot be unknown to anyone, in that they lie in reason for all; hence nobody is guiltless there in such error, but in regard to positive law there are *errores inculpabiles* [errors of fact], and there one may act in all innocence" (C 27:355).

Rivalry and emulation

The instinct of rivalry (also referred to as antagonism) involves the deep-seated inclination to compare oneself with others in regard to both moral merit and well-being. This instinct plays a role in the psychology of the devilish vices. Kant writes that the maxims of viciousness associated with these vices

[T]ake the ground of their origin from a property of human nature native to man, which not only makes us intrinsically guiltless, but also determines us to an admirable purpose: namely, the instinct [*Instinct*] of antagonism or rivalry, i.e., the inclination to work against the perfection of others, or to surpass them by ever-increasingly promoting our own cultivation, in agreement with the laws of morality. This is shown in appearance by the fact that we constantly compare ourselves with other men, and feel a chagrin on discovering their good points, whether it be their dutiful conduct, their honor, or their well-being. (V 27:692)

It is working against the perfection of others (their “good points”) that is characteristic of the other-regarding vices manifested, for example, in the vice of envy. In sharp contrast to this manifestation of rivalry, its manifestation in the cultivation of one’s own perfections in agreement with the laws of morality is what, we will see, Kant refers to as emulation (*Aemulation*). Notice that Kant’s talk of rivalry (at least in this passage) is something common to one’s working against the perfections of others, which leads to other-regarding vice, as well as to one’s engaging in a regimen of emulation. That is, rivalry as here understood need not be something nasty; there is such a thing as friendly rivalry even if it involves some discomfort or pain at viewing others as in some way superior to oneself. Aristotle makes this point nicely in comparing emulation with envy:

Emulation is pain caused by seeing the presence, in persons whose nature is like our own, of good things that are highly valued and are possible for ourselves to acquire; but it is felt not because others have these goods, but because we have not got them ourselves. It is therefore a good feeling felt by good persons, whereas envy is a bad feeling felt by bad persons. Emulation makes us take steps to secure the good things in question, envy makes us take steps to stop our neighbor from having them. (*Rhetoric*, Bk. II, II, 30–38 [McKeon 1941, 1402])

But notice also that in this passage from Vigilantius, Kant mentions the chagrin (something painful) that one feels in comparisons that lead to rivalry. Envy has as its aim, so to speak, the preservation of one’s sense of self-esteem – not only as an individual with equal standing to others but

also as an individual who craves appraisal esteem from others regardless of whether one is worthy of such esteem. It is the preservation of one's sense of self-esteem that drives the devilish vices.

However, it is important to take note of a distinction that is implicit in Kant's use of the term "emulation." In some places, Kant refers to emulation as an *impulse* (*Trieb*) implanted in human nature whose purpose "really lay in inciting men to constant cultivation of greater perfection in themselves, by comparison with others" (V 27:678f). It is thus something to be cultivated for this purpose. But he also notes this impulse can lead to the sort of rivalry in which one works against the perfections (and/or well-being) of others, a result of emulation he describes as "a side of human nature that has become malignant" (V 27:678). So emulation as an impulse is closely related to the instinct of rivalry. Here is how we understand their connection. Emulation necessarily involves comparison of oneself to others. This comparison triggers the instinct of rivalry in cases where one feels chagrin as a result of viewing the other party as having "goods" that one lacks. This rivalry can serve the purpose of emulation by taking the form of striving to be like the other person in some respect by engaging in some regimen with the aim of improving upon whatever it is that one sees as lacking in oneself and thus bringing oneself up to the other person's level, if not surpassing him or her. But the rivalry can also take a malignant form leading to one or another vice.

Rivalry, then, can serve a good purpose, namely spurring one to emulate one's rival by improving one's own condition, whether it be one's character or other advantages. Of course, pursuing a course of self-improvement does not mean that one is doing so out of a true love of honor. One might be striving to emulate one's rival because one is a mere craver of honor. However, the proper motivation guiding emulation should be one's love of honor, which, says Kant, is "imprinted" in every individual and is "the source and principle that finds expression in emulation – a drive constantly to perfect oneself in comparison with others" (V 27:680). Presumably, because love of honor (in a sense to be contrasted with love of honor as an acquired virtue) is "imprinted" it functions as a motivational source of emulation – an inclination (habitual desire) to be at least equal to others in those respects deserving of approval. Because love of honor (in contrast to a craving for outward undeserved recognition) has as its object proper self-esteem, and thus comparison with the moral law as a measure of that esteem, it serves as a principle guiding emulation.

In addition to the drive to be respected, the instinct of rivalry, and the impulse of emulation, there is one more ingredient essential to the dynamics of virtue in Kant's theory, to which we now turn.

Proper self-contentment

As remarked earlier, an important element in coming to be a lover of honor is having a proper sense of, and control over, one's happiness which, Kant says, is founded on "being contented with oneself" (V 27:649). The contentment in question requires not allowing one's needs to expand beyond what is really necessary for avoiding physical and psychological ills, and so taking care of one's needs as a physical being, and in particular not letting "amenities of life" become needs (see V 27:646–657, §§ 104 and 106). In these passages, Kant mentions a "principle of maximum dispensability" regarding life's amenities which is summed up by the rule: "Try to maintain your state so that you dispense with as much as is needed to establish your contentment on the fewest possible conditions" (V 6:649, see also C 27:392–406). What Kant is advocating in these passages is the importance of coming to adopt an attitude of self-contentment toward one's well-being or happiness that does not undermine one's aspiring to true love of honor, and so coming to settle on a conception of one's happiness that is within the bounds of morally practical reason. As we shall see in [Section 3](#), this ingredient is a crucial psychological element in avoidance of vice.

We are now in a position to characterize, at a general level of description, the psychology of virtue acquisition and maintenance, or at least the steady path of progress toward Kant's moral ideal of self-perfection. What we learn from the *Lectures*, as lately noted, is that there are certain natural drives and impulses whose proper functioning is important for successfully engaging in the project of self-perfection. The overall process in question is unquestionably social. To be a person of virtue, one needs the kind of moral understanding that the drive to be respected by others prompts, viz., consulting the moral judgments of (certain) others. Knowing one's duty is part of moral virtue (TL 6:446). But self-perfection and thus virtue also require that we engage in comparing ourselves with others – the result of the impulse to rivalry – whose manifestation as emulation guides one toward developing those traits that contribute to virtue. Specifically, through emulation in cases where one is focused on the moral qualities of one's rivals, one is motivated to acquire or develop more fully the moral virtues, but not because one is in competition with one's rivals. The proper basis for judging the quality of one's moral character is the ideal of humanity prescribed by the moral law. The role of select others in regard to one's pursuing virtue is that they serve to make manifest by example the kinds of behavior and attitudes that one should strive to acquire or more fully

develop. Emulating others for this purpose is an essential element in a human's pursuit of virtue.⁵

As noted earlier, the instinct to rivalry also involves comparison with others regarding one's own well-being. And unless one is able to be content in life without needing the amenities others may enjoy, including the urge to acquire material goods, or have certain powers over others, or whatever, the likely discontent that arises from not having such things will tend to disrupt the proper operation of emulation. Proper contentment with one's welfare requires so-called structural virtues – virtues that concern “types of strength in rational self-government.”⁶ These include virtues the possession of which serves not only to control one's inclinations so that they don't become needs, but also to gain mastery over one's emotions and passions. In places, Kant refers to the duty (and corresponding virtue) of self-mastery, characterized (qua duty) as a “positive command to a human being, namely to bring all his capacities and inclinations under his (reason's) control, and so to rule over himself” (TL 6:408).⁷ Frugality, “moderation or abstinence” in the possession of material goods is among the more specific virtues contributing to this kind of mastery (C 27:406, see also V 27:696). And this particular virtue has as one of its specific rules what we identified earlier as the *principle of maximum dispensability* – dispensing with “as much as is needed to establish your contentment on the fewest conditions possible” (V 6:649). So the elements that figure in the psychology of virtue acquisition and maintenance – its dynamic – involve the following:

- One needs an understanding of proper self-esteem (as an agent with dignity and as an imperfect being who is properly appraised for her actions and character), which concerns the moral law. *Recognition* of oneself as capable of moral law-giving and hence recognition of one's dignity, grounds recognitional self-esteem. *Comparison* of one's moral behavior and character with the demands of the moral law in regard to the ideal of humanity (a perfected human being), grounds appraisal self-esteem.
- Also needed are reliable judgments about the moral quality of one's behavior and attitudes that require comparing one's own views with those of others in order to guide one's progress – something the drive to be respected inclines one to do.

⁵ Kant claims that making such comparisons is a duty. See V 27:703f.

⁶ Adams 2006, 37. Following Roberts 1984, Adams contrasts structural virtues with motivational virtues.

⁷ At TL 6:408, Kant distinguishes the positive duty of self-mastery from the negative duty of apathy “forbidding him to let himself be governed by his feelings and inclinations.”

- Furthermore, comparison with others regarding moral merit and well-being for purposes of self-improvement is important. Such comparison calls upon the impulse of emulation, and in a good person results in one striving for those things of value (moral and nonmoral) that others have (or have to a greater extent), but that one currently lacks.
- In order for all of this to function properly, one must have a conception of well-being and a corresponding view of a contented life that is not in tension with the striving for self-perfection.
- Finally, in order for one's understanding of what constitutes a contented life to translate into behavior, one must possess the general structural virtue of self-mastery and the more specific virtues such as frugality that contribute to self-mastery.

When all of these elements are in good working order, one moves toward the ideal of moral perfection that constitutes virtue. One possesses the virtue of love of honor. But things can go wrong.

3 Emulation and the devilish vices

It is not entirely clear from the texts we have consulted what it is about envy, ingratitude, and *Schadenfreude* that makes them devilish. In Collins we find these remarks:

All three, ingratitude (*ingratitude qualificata*) [aggravated ingratitude, hatred of a benefactor], envy, and *Schadenfreude*, are devilish vices because they evince an immediate inclination to evil. That man should have a mediate inclination to evil is human and natural; the miser, for example, would like to acquire everything; but he takes no pleasure in the other having nothing at all. There are vices, therefore, that are both evil directly and indirectly. These three are those that are directly evil. (C 27:440)⁸

In this passage, the emphasis is on the inclination toward some form of other-directed evil for no further reason. This seems to be the sense in which the inclination to evil characteristic of devilishness is immediate. So, for instance, in the Herder notes we find this remark: "A direct pleasure at the other's misfortune is devilish" (H 27:5).⁹ The devilishness here apparently consists in taking underivative pleasure in the *state of affairs* of another's loss.

But notice the reference to hatred in this quote. It would seem that one could have an immediate inclination to take pleasure in another person's

⁸ The bracketed material is the footnote to the Latin phrase.

⁹ See also Kant's remark in Collins (C 27:439) about the difference between having a grudging attitude toward another's advantages and having the attitude of envy toward those advantages.

misfortune without hating the person. An episode of anger directed toward another person may cause one to take a passing immediate delight in that person undergoing some misfortune. Perhaps these vices as traits and their manifestations in thought and behavior need not include hatred to be devilish. One might suppose, then, that hatred just contributes to the *level* of devilishness manifested in certain instances of envy, ingratitude, and *Schadenfreude* as Kant maintains that the devilishness of these vices admits of degrees.

But, then again, perhaps hatred is essential to the possession of one of these vices counting as devilish. Anger is what Kant refers to as an affect, whereas hatred is a passion, and from the following quote it would seem that it is passion that is essential to vice if it is to be “properly” evil.

A passion is a sensible *desire* that has become a lasting inclination (e.g., *hatred*, as opposed to anger). The calm with which one gives oneself up to it permits reflection and allows the mind to form principles upon it and so, if inclination lights upon something contrary to the law, to brood upon it, to get it rooted deeply, and so to take up what is evil (as something premeditated) into its maxim. And the evil is then *properly* evil, that is, a true *vice*. (TL 6:408)

So, perhaps it is only when possession of envy, ingratitude, and *Schadenfreude* as traits and their corresponding episodes involve out and out hatred of the other that they are devilish.¹⁰ When speaking of their devilish nature, Kant describes such traits and corresponding episodes as directly contrary to a humane attitude toward others. Indeed, he remarks, because they “are directly opposed to moral sympathy, [they] “indicate *inhumanity*” (V 27:692). Sympathy, for Kant, is a sentiment, and it would seem that what is directly opposed to this sentiment is the sentiment (passion) of hatred. Moreover, Kant makes clear that not all episodes of these vices are devilish: it is only on “attaining their full height, [that] these are devilish vices” (C 27:439). These considerations strongly suggest, we think, that hatred is an essential element that contributes to making an otherwise nondevilish instance of a vice something devilish. The idea in connection with envy would be that an immediate inclination to take displeasure in another’s advantage, coupled with hatred of that person, is

¹⁰ In the *Tugendlehre*, envy as a character trait is characterized as involving “a propensity to view the well-being of others with distress, even though it does not detract from one’s own” (TL 6:458). Kant says of this propensity *to view* the well-being of others with distress that it is only an “indirectly malevolent disposition,” presumably because it does not (or need not) involve either an immediate inclination to negatively affect the well-being of the person one envies or hatred of the person envied.

what makes particular instances of envy devilish. In support of this interpretation, consider this remark about ingratitude:

If this ingratitude increases so much that he cannot endure his benefactor, and becomes his enemy, that is the devilish degree of vice, since it is utterly repugnant to human nature, to hate and prosecute those who have done one a kindness. (C 27:439)

Because devilishness admits of degrees, one could maintain that hatred is essential to devilish instances of these vices while also maintaining that the intensity of one's hatred affects their degree of devilishness.

Finally, there is another factor that pertains to the degree of viciousness (if not devilishness) of vice. When it comes to manifestations of the vices of envy, ingratitude, and *Schadenfreude* in thought and behavior, it is important to keep in mind the difference between the attitudes of wishing, wanting, intending (adopting a maxim), and acting on one's maxim. One can want the person one envies to meet with harm or misfortune without adopting a maxim of harming him and thus without forming the intention of doing so. And surely, adopting a maxim grounded in envy is worse than merely wanting some misfortune to befall the envied person. Furthermore, *acting* on such a maxim, which Kant refers to as envy *proper*, would seem to manifest a greater degree of viciousness than having the maxim but never acting on it (cf. TL 6:458). So there are degrees of envy that range from wishing or wanting the object of one's envy to be deprived of select advantages that one envies to more vicious manifestations of the vice. And, of course, the same may be said of ingratitude and *Schadenfreude*. If not all instances of wishing, wanting, intending, and acting that manifest one of these vices need be devilish, then one may draw a distinction between degrees of viciousness and degrees of devilishness, although perhaps, for example, if one's envious attitude leads one to engage in a harmful action out of envy (envy proper), it is very likely that one's envy involves those elements that make one's action devilish.

So with regard to Kant's devilish vices we have these questions. Given that not all instances of these vices (as traits) and their associated emotional episodes are devilish, what distinguishes the nondevilish instances from the devilish ones? Furthermore, because the devilishness of a vice is a matter of degree, which factor (or perhaps factors) affects their degree of devilishness? In response, here are our tentative proposals. First, we are inclined toward the view that hatred is an essential element in Kant's devilish vices being genuinely devilish. Second, one factor that affects the degree of devilishness manifested in these vices is the intensity of one's hatred. Perhaps another factor

(no doubt related to the first) is the attitude manifested in episodes of the vice and in particular whether one acts on one's hatred.¹¹

Let us now turn to the psychology of these vices, beginning with envy.

Envy

Episodes of envy, in Kant's view, involve at a minimum not wanting select others to have whatever "advantage" it is that those others have that one currently lacks.¹² "The envious man wishes to be happy when all around him are unhappy, and seeks the sweetness of happiness in this, that he alone enjoys it, and all other are unhappy" (C 27:438). How does envy come about?

Kant writes that "[e]nvy has its immediate ground in emulation," which, of course, refers to the *impulse* of emulation that impels us to compare ourselves with others. As we have noted, the purpose of this impulse is to spur one to cultivate those perfections that contribute to one's overall character and circumstances. And, as we know from Kant's emphasis on the epistemological import of self–other comparison, all goes well so long as one has an understanding of proper self-esteem, which involves comparing one's conduct and character with the standards of perfection as set forth by the moral law – the "method" that is characteristic of the lover of honor.

But [in the case of the devilish vices] this is where the error lies. That method [reflection on and comparison with the moral law] would establish true love of honor in us. But in judging himself and his inner worth, man founds and measures it instead on a merely comparative estimate of his person and condition against the worth and condition of other men. Hence arises the love of honor in a bad sense: if he finds himself lowered by comparison with the other, that arouses in him dislike of the other's person, and instead of actively exerting himself to become equal in value with the other, he succumbs to resentment at the latter's worth and merit, or tries to diminish him. (V 27:695)

So the impulse of emulation, together with a misunderstanding of proper self-esteem (and thus a failure to be a lover of true honor), are at the heart of

¹¹ Although, as noted, we allow that one may want to distinguish degrees of viciousness from degrees of devilishness.

¹² Merely misliking the fact that others have such advantages does not qualify as envy. Such misliking is a vice, but it is the "human" vice of begrudgement. And note that misliking the state of affairs in question is not the same as wanting that the others not have whatever advantage they happen to enjoy that one lacks. It is when this misliking escalates to a desire that others be deprived of those advantages that mere begrudgement escalates into envy. In some places Kant makes the point that what is particularly devilish about this vice is that "We say of begrudgement that a man can endure no one happy *above him*, but of envy that he can endure no one happy *beside him*" (V 27:698).

the psychology of envy. We can bring the various elements of the psychology of this vice into focus by setting forth the various stages leading to its full devilish manifestation.

- *Emulation-comparison stage*: Prompted by the impulse of emulation, one compares oneself to others with regard to characteristics that one particularly cares about because one takes them to bear on one's self-esteem. In this stage, one perceives self-other inequalities of one or another sort.
- *Rivalry stage*: Because one is focused on advantages the other has that one lacks, this prompts a feeling of *chagrin*, which triggers the instinct of rivalry.
- *Evaluation-decision stage*: The chagrin one feels (something unpleasant) is an emotional reaction to a perceived threat to one's self-esteem because one bases one's self-esteem entirely on the comparison with others (rather than the moral law) and thus "finds himself lowered by comparison with the other" (V 27:695). This evaluation prompts one not only to dislike the fact that there is someone who is "above" oneself (begrudgement), but this painful evaluation, together with other contributing factors (factors having to do with not being inoculated against vice), leads to a decision that "in order to get equal with him [one's rival] . . . [one will aim at] "the destruction of the other's well-being, to make him less fortunate than ourselves" (V 27:693). When one's other-directed hostility turns into hatred of one's rival, one's envy is devilish.
- *Other-directed hostility stage*: This decision leads one to engage in other-directed hostilities that are characteristic of what Kant refers to as "envy proper."

It is the evaluation-decision stage where the error lies.

However, these stages do not reveal details about the specific manner in which one can succumb to the vice of envy. Given that there are two conceptions of self-esteem (recognitional and appraisal), both of which are properly grounded in understanding the significance of and comparing oneself with the moral law, and given that, broadly speaking, one can be envious of another's moral standing or another's well-being, or both, there are various specific ways in which this vice can manifest. Here are a few examples.

One route starts from deficient recognitional self-esteem – the self-esteem one ought to have in virtue of being autonomous, something one shares with all others regardless of one's conduct or character and so in virtue of which one enjoys equal moral standing with all others. If one feels lowered in moral standing compared to others because of their nonmoral amenities, then one has a false view of moral standing: all are on equal

footing given autonomy. The amenities of others are not a relevant basis of comparison. The same can be said for feeling lowered because of another's moral merits.

If one feels lowered in appraisal self-esteem because of the other's amenities, then one likely fails to have a proper sense of nonmoral self-contentment; one has allowed one's conception of true happiness to be hostage to things, talents, or positions that are not necessities in life. And, in any case, proper appraisal self-esteem has to do with fulfilling moral requirements and performing morally meritorious actions. Of course, one may not be envious of others because they enjoy all sorts of amenities in life; one may envy them rather because they are not in the sort of unfortunate circumstances one finds oneself in. Regarding this sort of situation, in Collins we find this remark about envy and one's fortunes in life: "To keep the soul free from the vice of envy, we must therefore try to bear every hardship, and, once it has befallen us, to extract from it the advantage that always resides in misfortune" (C 27:368).

If one feels lowered in appraisal self-esteem because of the other's superior moral merits, it is fantasy to suppose that one can restore one's self-esteem and thus preserve a true sense of honor by doing something prejudicial to their moral standing that lowers them. Lowering the other person does not restore or preserve proper self-esteem. Again, proper appraisal self-esteem is something one earns by complying with or exceeding the strict demands of the moral law.

Ingratitude

The psychological profile of this vice follows the same pattern as the one for envy.¹³ The error occurs at the evaluation-decision stage. The essential difference, of course, is that one feels lowered by one's benefactor as a result of the beneficiary–benefactor relationship. Ingratitude involves "a displeasure or discontent at the obligation the other has laid on us, through the kindness he has shown towards us" (V 27:694f). Again, the source of this vice is a faulty conception of one's self-esteem as Kant makes clear in this passage:

The fact that the other has done more for us than he was required to do, arouses ingratitude; for all his *merita* in regard to our person, or fortunes, bring it about that we are thereby obliged to him on that account; but in that he has had an influence on our well-being, he has an advantage over us,

¹³ We discuss the psychology of ingratitude in Smit and Timmons 2011.

whereby he is elevated above our worth, and we, on the contrary, have become *inferiores* in his regard; *for assuming that the estimation of our self rests on a comparative judgment with the other's worth*, this degradation displeases us; we are tortured by the obligation he has laid upon us, and so envy hinders our participation and interest in his welfare. (V 27:695f, our emphasis immediately after the semicolon)

Again, Kant is clear in the Collins notes that this vice need not be devilish. In considering the question of whether “the human soul contains an immediate inclination to evil, and thus a propensity for devilish vice,” Kant wonders whether a person can “be so ungrateful as actually to hate his benefactor” as episodes of ingratitude may only involve being “far too proud to be thankful to [one's benefactor], and for the rest, wishes him every happiness; the only thing is, he would like to be well out of his way” (C 27:440f). Such episodes of ingratitude are not devilish, although morally objectionable. With regard to devilish episodes (and the associated trait of ingratitude), there is room, as explained earlier, to admit degrees of devilishness.

Schadenfreude

Kant claims that all three devilish vices “take their rise from an inborn tendency to set oneself against the other's superiority” (V 27:695); rivalry in other words. So like envy and ingratitude, typical manifestations of *Schadenfreude* involve self–other comparison. Space does not permit us to delve into the psychology of *Schadenfreude* as Kant sees it; we confine ourselves to a few remarks.

In *Vigilantius* Kant associates this vice with envy. “*Schadenfreude (vugaris)* is malevolence or joy at another's misfortune, which seems, therefore, to be founded on envy, or discontent at his happiness” (V 27:695). However, the “seems” in this passage casts some doubt on whether Kant is saying that *Schadenfreude* really is grounded on envy. Moreover, it is not clear what talk of “grounding” means in this passage. Does it mean that a general explanation of its psychological genesis will appeal to the vice of envy? This seems unlikely. Nowhere in the *Lectures on Ethics* or in Kant's published work do we find this claim defended. Moreover, in Collins (C 27:441) Kant claims that this vice is often “already strongly apparent in the young” when they engage in certain forms of mischief. He remarks that this inclination to take joy in another's distress is “a sort of animality, whereby man retains something of the beast in him, which he cannot overcome. The source of it we know not, and for some of our characteristics we can adduce no reason

whatsoever.” There is no mention of envy here.¹⁴ What does seem plausible is that one who is envious of another’s merits, either in wish or desire, will very likely take an immediate delight in the object of their envy losing the merits one envies. *Schadenfreude* is a natural companion of envy, but one can take joy in others’ misfortunes without first envying them.

So let us put this passage aside and focus on Kant’s remark about the devilish vices arising from an inborn tendency to be opposed to superiority in others. Here we turn for a moment to passages in the *Tugendlehre*. In discussing the vice of malice – “the direct opposite of sympathy” (TL 6:459), which involves *Schadenfreude* at its core – Kant remarks “It is, indeed natural that, by the laws of imagination (namely, the law of contrast), we feel our own well-being and even our good conduct more strongly when the misfortune of others or their downfall in scandal is put next to our own condition, as a foil to show it in so much the brighter light” (TL 6:460). But how does this focus on one’s own comparative well-being or moral merit transform into taking *immediate* pleasure in the misfortunes of others, where the focus of concern has shifted from self to others? In this same passage, Kant seems to be addressing this question:

It is the *haughtiness* of others when their welfare is uninterrupted, and their *self-conceit* in their good conduct (strictly speaking, only in their good fortune in having so far escaped temptations to public vice) – both of which an egotist accounts to his merit – that generate this malevolent joy. (TL 6:460)

These remarks suggest that at least one typical source of *Schadenfreude* is a response to what one judges to be the unjustified presumption of superiority on the part of others. One’s natural resentment of such people likely triggers an immediate dislike of them, and perhaps justifiably so. But Kant distinguishes dislike from ill-will, the latter characterized as a hatred of others that includes the vice of *Schadenfreude* (C 27:430–432). If one focuses on the remarks about haughtiness and self-conceit, then one can see how our psychological profile of envy and ingratitude also fits the vice of malicious joy. *Emulation-comparison stage*: the impulse of emulation prompts one to compare oneself with others, particularly as it bears on one’s self-esteem. *Rivalry stage*: the display of such traits as haughtiness and self-conceit in others, which one naturally interprets as intended by

¹⁴ However, some theorists of emotion distinguish primitive from sophisticated forms of various emotions, including envy, and so it may be that a primitive form of envy is the basis of the kind of mischief to which Kant is referring. On the primitive/sophisticated distinction regarding envy and jealousy, see Taylor 1988.

them to affirm their superiority over oneself, triggers the instinct of rivalry. *Evaluation-decision stage*: because one's self-esteem is based entirely on comparison with others, their attitude of superiority results in one's coming to have ill-will toward them, which includes a disposition to take malicious joy in any misfortunes they suffer. *Other-directed hostility stage*: one's ill-will leads one to engage in other-directed hostilities and thus what Kant calls "malice proper" (TL 6:460).

There is no mention of envy in this story, and, indeed, one need not be envious of those who display haughtiness or objectionable self-conceit. Is this psychological profile general enough to cover all cases of this vice? Certainly episodes of this vice are not only directed toward the arrogant, and it seems implausible that *Schadenfreude* always begins with hostility toward the arrogant and then takes root in one's personality. There is obviously more interpretive work to do in exploring the psychology of this vice in Kant's thought. But we leave this work for another occasion.

Finally, like the two other devilish vices, there are episodes of genuine *Schadenfreude* that are not devilish. Kant notes that "*Schadenfreude* . . . is not always rooted in wickedness . . . [because] it often has its source in mere *mischief*; glee at another's mishap" (V 27:698). He gives a few examples, including "glee at the other's antics when he falls, the mischievous frightening of others to enjoy their alarm." He claims that "there is in all these cases an ethically objectionable *Schadenfreude*" (ibid.), despite their not being devilish, since glee at another's mishaps need not evince an immediate inclination to evil that expresses other-directed hatred. Regarding devilish episodes of this vice, they come in degrees subject to the same factors that determine the degrees of devilishness manifested in episodes of envy and ingratitude.

4 Conclusion

We announced at the outset that in addition to providing an interpretation of the psychology of the devilish vices in Kant, we are interested in the extent to which this cluster of vices is psychologically unified. This unification issue itself splits into two questions. First, whether, and to what extent, the genesis of these vices as character traits is the same, and second, whether, and to what extent, particular episodes (at a general level of description) have the same or similar psychological profile. Here are our somewhat tentative answers to these questions. Kant says that these three vices *arise* from our natural tendency to oppose others' superiority, and our psychological account of each of these vices captures this idea. Again,

episodes of envy and ingratitude, we have argued, share the same psychological pattern, a pattern that is also evident in at least typical cases of *Schadenfreude*. In addition, there is a common ingredient in all truly devilish manifestations of these vices, namely the element of other-directed hatred, which is opposed, and not merely indifferent to, the global obligatory end of promoting the well-being of others.

Works cited

- Achenwall, Gottfried 1763. *Prolegomena Iuris Naturalis* (2nd ed.) Göttingen: Bossigel
1767. *Prolegomena Iuris Naturalis* (3rd ed.) Göttingen: Bossigel
- Achenwall, Gottfried and Johann Stephan Pütter 1995. *Anfangsgründe des Naturrechts (Elementa Iuris Naturae)*, ed. and trans. Jan Schröder (ed., trans.). Frankfurt am Main: Insel Verlag
- Adams, Robert M. 2006. *A Theory of Virtue: Excellence in Being for the Good*. Oxford University Press
- Aichele, Alexander 2004. “Sive vox naturae sive vox rationis sive vox Dei? Die metaphysische Begründung des Naturrechtsprinzips bei Heinrich Köhler mit einer abschließenden Bemerkung zu Alexander Gottlieb Baumgarten,” *Jahrbuch für Recht und Ethik/Annual Review of Law and Ethics* 12: 115–135
2005. “Die Ungewißheit des Gewissens: Alexander Gottlieb Baumgartens forensische Aufklärung der Aufklärungsethik,” *Jahrbuch für Recht und Ethik/Annual Review of Law and Ethics* 13: 3–30
- Allison, Henry 1990. *Kant’s Theory of Freedom*. Cambridge University Press
- Ameriks, Karl 2012. *Kant’s Elliptical Path*. Oxford: Clarendon Press
- Ameriks, Karl and Steve Naragon (trans., ed.) 1997a. Immanuel Kant, *Lectures on Metaphysics*. Cambridge University Press
- 1997b. “Translators’ Introduction” to Ameriks and Naragon 1997a, xiii–xliii
- Annas, Julia 1993. *The Morality of Happiness*. Oxford University Press
2006. “Virtue Ethics,” in *The Oxford Handbook of Ethical Theory*, ed. David Copp. Oxford University Press, 515–536
- Aristotle. *Politics* in Barnes (ed.) 1984, 1986–2129
- Rhetoric* in McKeon (ed.) 1941, 1317–1451
- Arnoldt, Emil 1909. *Gesammelte Schriften*, vol. 5. Otto Schöndörffer (ed.) Berlin: Bruno Cassirer
- Bacin, Stefano 2006. *Il senso dell’etica. Kant e la costruzione di una teoria morale*. Bologna: il Mulino
2013. “The Perfect Duty to Oneself Merely as a Moral Being,” in Trampota, Sensen, and Timmermann 2013, 159–181
- Bacin, S., A. Ferrarin, C. La Rocca, and M. Ruffing (eds.) 2013. *Kant und die Philosophie in weltbürgerlicher Absicht: Akten des XI. Internationalen Kant-Kongresses 2010*. Berlin: Walter de Gruyter

- Barnes, Jonathan (ed.) 1984. *The Complete Works of Aristotle*. 2 vols. Princeton University Press (*Politics*, 1986–2129)
- Baron, Marcia 1995. *Kantian Ethics Almost without Apology*. Ithaca, NY: Cornell University Press
- Baumgarten, Alexander Gottlieb 1739. *Metaphysica*. Editio tertia, Halle 1757. [Reprinted in *Kants Gesammelte Schriften*. Berlin: Walter de Gruyter AA 17: 5–226 and 15: 5–54. English translation: Fugate and Hymers 2013]
- 1740 (1st ed.), 1751 (2nd ed.), 1763 (3rd ed.) *Ethica Philosophica*. Halle [2nd and 3rd eds. reprinted in AA 27: 737–869, 27: 873–1028]
1760. *Initia philosophiae practicae primae*. Halle [Reprinted in AA 19: 7–91]
1761. *Acroasis logica*. In *Christianum L. B. de Wolff*. Halle [Reprint: Hildesheim: Olms, 1973 (Wolff, *Gesammelte Werke*, Abt. III, Bd. 5)]
1770. *Philosophia generalis, edidit cum dissertatione proemiali de dubitatione et certitudine Ioh. Christian Foerster*, Halle [Reprint: Hildesheim: Olms, 1968]
- Baumgarten, Siegmund Jacob 1738. *Unterricht vom rechtmäßigen Verhalten eines Christen oder Theologische Moral*. Halle
- Baxley, Anne Margaret 2010. *Kant's Theory of Virtue: The Value of Autocracy*. Cambridge University Press
- Beck, Lewis White 1960. *A Commentary on Kant's Critique of Practical Reason*. University of Chicago Press
1963. "Foreword" to Immanuel Kant, *Lectures on Ethics*, Louis Infield (trans.). New York: Harper & Row, ix–xiv
- Beiser, Frederick 2005. *Schiller as Philosopher: A Reexamination*. Oxford University Press
- Betzler, Monica (ed.) 2008. *Kant's Theory of Virtue*. Berlin: Walter de Gruyter
- Blackburn, Simon 1993. *Essays in Quasi-Realism*. Oxford University Press
- Brewer, Talbot 2002. "The Character of Temptation: Towards a More Plausible Kantian Moral Psychology," *Pacific Philosophical Quarterly* 83: 103–130
- Brink, David 1999. "Review of S. Engstrom and J. Whiting (eds.), *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*," *Philosophical Review* 108: 576–582
2000. "Review of Nancy Sherman, *Making a Necessity of Virtue: Aristotle and Kant on Virtue*," *Philosophical Review* 109: 428–434
- Brandt, Reinhard. 1982 "Das Erlaubnisgesetz, oder: Vernunft und Geschichte in Kants Rechtslehre," in *Rechtsphilosophie der Aufklärung*, R. Brandt (ed.), Berlin: Walter de Gruyter, 233–285
2004. "Das Problem der Erlaubnisgesetze im Spätwerk Kants," in *Zum Ewigen Frieden*, Otfried Höffe (ed.), Berlin: Akademie Verlag, 69–86
- Butler, Joseph 2002 [1729]. *Fifteen Sermons and the Dissertation on the Nature of Virtue*. Cambridge University Press
- Byrd, B. Sharon 1989. "Kant's Theory of Punishment: Deterrence in its Threat, Retribution in its Execution," *Law and Philosophy* 8: 151–200
2010. "Intelligible Possession of Objects of Choice," in Denis 2010a, 93–110
- Byrd, B. Sharon and Joachim Hruschka 2010. *Kant's Doctrine of Right: A Commentary*. Cambridge University Press

- Carnois, Bernard 1987. *The Coherence of Kant's Doctrine of Freedom*, David Booth (trans.) University of Chicago Press
- Carmichael, Gershom 1739 [1718, 1724] *Observationes et Supplementa ad Cl. Puffendorffii libros duos De Officio Hominis et Civis*, reprint of 2nd ed. Basel: Brandmüller
- Cooper, John M. (ed.) 1997. *Plato: Complete Works*. Indianapolis: Hackett. (*Euthydemus*, 708–745; *Meno*, 870–897; *Republic*, 971–1223; *Laws*, 1318–1616)
- Crisp, Roger and Michael Slote (eds.) 1997. *Virtue Ethics*. Oxford University Press
- Crusius, Christian August 1969 [1744]. *Anweisung vernünftig zu leben*, Leipzig [Reprint ed. by G. Tonelli, Hildesheim: Olms]
- Daries, Joachim Georg 1740. *Institutiones Iurisprudentiae Universalis*. Jena: Buchius
1754. *Observationes Iuris Naturalis, Socialis et Gentium*. Jena: Güth
- Delfosse, Heinrich P., Norbert Hinske and Gianluca Sadun Bordoni (eds.) 2010. *Stellenindex und Konkordanz zum "Naturrecht Feyerabend"*, vol. 1: *Einleitung des "Naturrechts Feyerabend"*. Stuttgart-Bad Canstatt: Fromann-Holzboog
- Denis, Lara 1999. "Kant on the Wrongness of 'Unnatural' Sex," *History of Philosophy Quarterly* 16 (2): 225–248
- (ed.) 2010a. *Kant's Metaphysics of Morals: A Critical Guide*. Cambridge University Press
- 2010b. "Freedom, Primacy, and Perfect Duties to Oneself," in Denis 2010a, 170–191
- 2010c. "Humanity, Obligation, and the Good Will: An Argument against Dean's Interpretation of Humanity," *Kantian Review* 15: 118–141
2014. "Love of Honor as a Kantian Virtue," in *Kant on Emotion and Value*, Alix Cohen (ed.). Hampshire, UK: Palgrave Macmillan, 191–209
- Dyck, Corey W. 2012. "Chimerical Ethics and Flattering Moralists: Baumgarten's Influence on Kant's Moral Theory in the *Observations and Remarks*," in Shell and Velkley 2012, 38–56
- Ebels-Duggan, Kyla 2008. "Against Beneficence: A Normative Account of Love," *Ethics* 119: 142–179
- Engisch, Karl 1960. *Logische Studien zur Gesetzesanwendung* (2nd ed.) Heidelberg: Winter
- Engstrom, Stephen 1997. "Kant's Conception of Practical Wisdom," *Kant-Studien* 88: 16–43
2002. "The Inner Freedom of Virtue," in *Kant's Metaphysics of Morals: Interpretive Essays*, Mark Timmons (ed.). Oxford University Press, 289–316
- Feinberg, Joel 2008. "Psychological Egoism," in *Reason & Responsibility*, Feinberg, Joel and Shafer-Landau, Russ (eds.). Belmont, CA: Thomson Wadsworth, 520–532
- Ferrarin, Alfredo 2013. "The Unity of Reason: On Cyclopes, Architects, and the Cosmic Philosopher's Vision," in Bacin, Ferrarin, La Rocca and Ruffing 2013, vol. 1, 213–228
- Feuerbach, Ludwig 1971 [1841]. *Das Wesen des Christentums*. Stuttgart: Reclam
- Fichte, Johann Gottlieb 2005 [1798]. *Das System der Sittenlehre*. [English translation ed. by Daniel Breazeale and Günter Zöller. Cambridge University Press]

- Flikschuh, Katrin 1999. "Freedom and Constraint in Kant's *Metaphysical Elements of Justice*," *History of Political Thought* 20: 250–271
2000. *Kant and Modern Political Philosophy*. Cambridge University Press
2004. "Ist das rechtliche Postulat ein Postulat der reinen praktischen Vernunft? Zum Endzweck der Kantischen Begriffe des Rechts," *Jahrbuch für Recht und Ethik/Annual Review of Law and Ethics* 12: 299–330
- Frierson, Patrick 2011. "Introduction" to Immanuel Kant, *Observations on the Feeling of the Beautiful and Sublime and Other Writings*, Patrick Frierson and Paul Guyer (eds.) Cambridge University Press, vii–xxxv
2012. "Two Concepts of Universality in Kant's Early Ethics," in Shell and Velkley 2012, 57–76
- Fugate, Courtney D. and John Hymers (trans., ed.) 2013. Alexander Baumgarten, *Metaphysics: A Critical Translation with Kant's Elucidations, Selected Notes, and Related Materials*. London: Bloomsbury
- Gawlick, Günter and Lothar Kreimendahl 2011. "Einleitung," to Alexander Gottlieb Baumgarten, *Metaphysica. Metaphysik. Historisch-kritische Ausgabe*. Stuttgart-Bad Cannstatt: Frommann-Holzboog, ix–lxxxvii
- Gottsched, Johann Christoph 1983 [1733–1734]. *Erste Gründe der gesamten Weltweisheit*. Leipzig: Breitkopf, 1762 [Reprint: Hildesheim: Olms, 2 vols. (Wolff, *Gesammelte Werke*, III Abt., Bd. 20.1–2)]
- Grenberg, Jeanine 2005. *Kant and the Ethics of Humility: A Story of Dependence, Corruption, and Virtue*. Cambridge University Press
- Guyer, Paul 1993. *Kant and the Experience of Freedom: Essays on Aesthetics and Morality*. Cambridge University Press
2005. *Kant's System of Nature and Freedom: Selected Essays*. Oxford University Press
2006. *Kant*. London: Routledge
2010. "Moral feelings in the *Metaphysics of Morals*," in Denis 2010a, 130–151
2011. "Kantian Perfectionism," in *Perfecting Virtue. New Essays on Kantian Ethics and Virtue Ethics*, Lawrence Jost and Julian Wuerth (eds.) Cambridge University Press, 194–214
- Groß, Felix (ed.) 1912. *Immanuel Kant. Sein Leben in Darstellungen von Zeitgenossen (Die Biographien von L.E. Borowski, R. B. Jachmann und A. Ch. Wasianski)*. Berlin: Deutsche Bibliothek
- Hare, John E. 1996. *The Moral Gap: Kantian Ethics, Human Limits and God's Assistance*. Oxford University Press
2009. "Is Moral Goodness without Belief in God Rationally Stable?" in *Is Goodness without God Good Enough?* Robert K. Garcia and Nathan L. King (eds.) Lanham, MD: Rowman & Littlefield, 85–99
- Heath, Peter (trans., ed.) and J. B. Schneewind (eds.) 1997. Immanuel Kant, *Lectures on Ethics*, Cambridge University Press
- Hegel, G. W. F. 1991 [1827]. *The Encyclopedia Logic*. T. F. Geraets, W. A. Suchting and H. S. Harris (trans.) Indianapolis: Hackett
- Heineccius, Johann Gottlieb 1738. *Elementa iuris naturae et gentium*. Halle [English translation: *A Methodical System of Universal Law*. Translated from the Latin

- by George Turnbull. Thomas Ahnert and Peter Schröder (eds.) Indianapolis: Liberty Fund, 2008]
- Herman, Barbara 1993. "Could It Be Worth Thinking About Kant on Sex and Marriage?" in *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, Louise M. Antony and Charlotte Witt (eds.). Boulder: Westview, 49–67
- Heumann, Herman Gottlieb 1926 [1907]. *Heumanns Handlexikon zu den Quellen des römischen Rechts*, 2nd reprint of 9th ed., reworked by E. Seckel, Jena: Fischer
- Höffe, Otfried 2010. "Kant's Innate Right as a Rational Criterion for Human Rights," in Denis 2010a, 71–92
- Hruschka, Joachim 1986. *Das deontologische Sechseck bei Gottfried Achenwall im Jahre 1767*. Hamburg: Joachim Jungius-Gesellschaft
1990. "The Hexagonal System of Deontic Concepts according to Achenwall and Kant," in *Estudios de Historia de la Logica, Actas del II Simposio de Historia de la Logica*, I. Angelelli and A. d'Ors (eds.). Pamplona: Universidad de Navarra, 277–294
1991. "Rechtfertigungs- und Entschuldigungsgründe: Das Brett des Karneades bei Gentz und bei Kant," *Goldammer's Archiv für Strafrecht*, 1–10
2001. "Die species facti und der Zirkel bei der Konstitution des Rechtsfalles in der Methodenlehre des 18. Jahrhunderts," in *Theorie der Interpretation vom Humanismus bis zur Romantik*, Jan Schröder (ed.). Stuttgart: Steiner, 203–214
2003. "Der Einfluß des Aristoteles und der Aristoteles-Rezeptionen auf die Bildung heutiger Rechtsbegriffe am Beispiel der 'actio libera in causa'," in *Bürgerliche Freiheit und Christliche Verantwortung*, Heinrich de Wall and Michael Germann (eds.). Tübingen: Mohr-Siebeck, 687–704
2004. "The Permissive Law of Practical Reason in Kant's *Metaphysics of Morals*," *Law and Philosophy* 23: 45–72
- Hruschka, Joachim and Jan C. Joerden 1987. "Supererogation: Vom deontologischen Sechseck zum deontologischen Zehneck – Zugleich ein Beitrag zur strafrechtlichen Grundlagenforschung," *Archiv für Rechts- und Sozialphilosophie* 73: 104–120
- Hume, David 1956 [1757]. *Natural History of Religion*. Stanford, CA: Stanford University Press
- 1967 [1739–40]. *A Treatise on Human Nature*, L. A. Selby-Bigge (ed.) Oxford: Clarendon Press
- 1978 [1739–40]. *A Treatise of Human Nature*, L.A. Selby-Bigge and P.H. Nidditch (eds.). Oxford University Press, 2nd ed.
- Hursthouse, Rosalind 1999. *On Virtue Ethics*. Oxford University Press
- Infield, Louis (ed., trans.) 1930. Immanuel Kant, *Lectures on Ethics*. London: Methuen [Reprinted: Indianapolis: Hackett, 1981]
- Irwin, Terence 2004. "Kantian Autonomy," *Royal Institute of Philosophy Supplement* 55: 137–164
2009. *The Development of Ethics. A Historical and Critical Survey*, vol. III: *From Kant to Rawls*. Oxford University Press
- Jachmann, Reinhold Bernhard 1804. *Immanuel Kant geschildert in Briefen an einen Freund*. Königsberg: F. Nicolovius

1993. "Immanuel Kant geschildert in Briefen an einen Freund" in *Immanuel Kant*, Felix Gross and Rudolf Malter (eds.) Darmstadt: Wissenschaftliche Buchgesellschaft, 103–187
- Johnson, Robert N. 2011. *Self-Improvement: An Essay in Kantian Ethics*. Oxford University Press
- Kain, Patrick 2004. "Self-Legislation in Kant's Moral Philosophy," *Archiv für Geschichte der Philosophie* 86: 257–306
2010. "Duties Regarding Animals," in Denis 2010a, 210–233
- Kalinowski, Georges 1972. *La logique des normes*. Paris: Presses universitaires de France
- Kaufmann, Matthias 2005. "Was erlaubt das Erlaubnisgesetz – und wozu braucht es Kant?" *Jahrbuch für Recht und Ethik/Annual Review of Law and Ethics* 1: 195–219
- Koehler, Heinrich 1738. *Juris naturalis eiusque cumprimis cogentis methodo systematica propositi Exercitationes VII*. Frankfurt
- König, Peter 1994. *Autonomie und Autokratie: Über Kant's Metaphysik der Sitten*. Berlin: Walter de Gruyter
- Korsgaard, Christine M. 2011. "Interacting with Animals: A Kantian Account," in *The Oxford Handbook of Animal Ethics*, Tom L. Beauchamp and R. G. Frey (eds.) Oxford University Press, 91–118
- Krauß, Wilhelm 1926. *Untersuchungen zu Kants moralphilosophischen Vorlesungen*. (PhD dissertation, University of Tübingen)
- Kuehn, Manfred 2001. *Kant: A Biography*. Cambridge University Press
2004. "Introduction" to Stark 2004a, vii–xxxv
- 2009a. "Ethics and Anthropology in the Development of Kant's Moral Philosophy," in Timmermann 2009a, 7–28
- 2009b. "Reason as a Species Characteristic," in *Essays on Kant's Universal History*, Amelie Rorty and James Schmidt (eds.). Cambridge University Press, 68–93
2010. "Kant's *Metaphysics of Morals*: The History and Significance of Its Deferral," in Denis 2010a, 9–27
2011. "Kant on Education, Anthropology and Ethics," in *Kant and Education: Interpretations and Commentary*, Klas Roth and Chris W. Surprenant (eds.). London: Routledge, 55–68
- Lehmann, Gerhard 1979. "Einleitung" to Immanuel Kant, *Vorlesungen über Moralphilosophie in Kants gesammelte Schriften*, vol 27. Berlin: Walter de Gruyter, 1037–1068 [AA 27]
- Louden, Robert B. 1992. *Morality and Moral Theory*. Oxford University Press
1997. "Some Vices of Virtue Ethics," in Crisp and Slote 1997, 201–216
2000. *Kant's Impure Ethics: From Rational Beings to Human Beings*. Oxford University Press
- 2007a. "Kantian Moral Humility: Between Aristotle and Paul," *Philosophy and Phenomenological Research* 75 (3): 632–639
- 2007b. *The World We Want: How and Why the Ideals of the Enlightenment Still Elude Us*. Oxford University Press
2011. *Kant's Human Being: Essays on His Theory of Human Nature*. Oxford University Press

- in press. "‘The End of All Human Action’/‘The Final Object of All My Conduct’: Aristotle and Kant on the Highest Good," in *The Highest Good in Aristotle and Kant*, Joachim Aufderheide and Ralf Bader (eds.). Oxford University Press
- Ludwig, Bernd (ed.) 1986, 1998. Immanuel Kant, *Metaphysische Anfangsgründe der Rechtslehre*. Hamburg: Felix Meiner
1988. *Kants Rechtslehre. Mit einer Untersuchung zur Drucklegung Kantischer Schriften von Werner Stark*. Hamburg: Felix Meiner
- MacIntyre, Alasdair 1984 (2nd ed.). *After Virtue*. University of Notre Dame Press
- Macmurray, John 1930. "Introduction" to Immanuel Kant, *Lectures on Ethics*, Louis Infield (trans.) New York: Harper & Row, xv–xix
- Martini, Karl Anton von 1770. *De Lege Naturali Exercitationes sex*. Vindobonae
- McKeon, R. (ed.) 1941. *The Basic Works of Aristotle*. New York: Random House (*Rhetoric*, W. Rhys Roberts [trans.], 1317–1451)
- Meier, Georg Friedrich 1763. *Alexander Gottlieb Baumgartens Leben*. Halle 1763 [Reprinted in *Aufklärung* 20 (2008), 351–373]
- 2008 [1764]. *Allgemeine practische Weltweisheit*, Halle 1764 [Reprint: Hildesheim: Olms]
- Mendelssohn, Moses 1767. *Phädon, oder über die Unsterblichkeit der Seele*. Berlin
- Menzer, Paul (ed.) 1924a. *Eine Vorlesung Kants über Ethik*. Berlin: Pan Verlag Rolf Heise
- 1924b. "Anmerkungen" to Menzer 1924a, 323–335
- 1991, "Einleitung (1924)," in Immanuel Kant, *Eine Vorlesung über Ethik*. Gerd Gerhard (ed.) Frankfurt a. Main: Fischer Taschenbuch Verlag, 281–289
- Merle, Jean-Christophe 1998. "Tugendlehre," in *Historisches Wörterbuch der Philosophie*, J. Ritter and K. Gründer (eds.), vol. 10. Basel: Schwabe, 1570–1572
- Naragon, Steve 2009. "Moral Philosophy Notes Vigilantius," <http://www.manchester.edu/kant/notes/notesMoral.htm#Vigilantius4>, visited on July 1, 2013
- 2010a. "Kant in the Classroom," <http://www.manchester.edu/Kant/Home/Index.htm>, visited on January 27, 2014
- 2010b. "Introduction: Kant's Lectures," <http://www.manchester.edu/kant/Lectures/lecturesIntro.htm>, visited on January 27, 2014
2014. "Moral Philosophy Notes," <http://www.manchester.edu/kant/Notes/notesMoral.htm>, visited on January 27, 2014
- Nichols, Shaun 2004. *Sentimental Rules*. Oxford University Press
- Nussbaum, Martha 2001. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge University Press
- Oberhausen, Michael and Riccardo Pozzo (eds.) 1999. *Vorlesungsverzeichnisse der Universität Königsberg (1720–1804)*. Stuttgart-Bad Cannstatt: Frommann-Holzboog
- O'Neill [Neill], Onora 1975. *Acting on Principle: An Essay on Kantian Ethics*. New York: Columbia University Press
- Palmquist, Stephen R. 2000. *Kant's Critical Religion*. Hampshire, UK: Ashgate Publishing

- Plato. *Euthydemus*, in Cooper 1997, 708–745
Meno, in Cooper 1997, 870–897
Laws, in Cooper, 1997, 1318–1616
Republic, in Cooper, 1997, 971–1223
- Pufendorf, Samuel 1989 [1672]. *Gesammelte Werke: De Jure Naturae et Gentium*. Frank Böhling (ed.) Berlin: Akademie Verlag
 1991 [1673]. *On the Duty of Man and Citizen According to Natural Law*. Michael Silverthornre (trans.) and James Tully (ed.). Cambridge University Press
 1997 [1673] *De Officio Homini et Civis*. Gerald Hartung (ed.). Berlin: Akademie Verlag
 1999 [1660], *Elementa jurisprudentiae universalis I*. Thomas Behme (ed.) Berlin: Akademie Verlag
- Rawls, John 1999. *A Theory of Justice: Revised Edition*. Cambridge, MA: Belknap Press
- Reath, Andrews 2006. *Agency and Autonomy: Selected Essays*. Oxford University Press
- Reich, Klaus 1939. “Kant and Greek Ethics” (Part I), *Mind* 48: 338–354
 2001. *Gesammelte Schriften*. Manfred Baum (ed.). Hamburg: Meiner
- Ricken, Friedo 2013. “Die Religionslehre als Lehre der Pflichten gegen Gott liegt außerhalb der Grenzen der reinen Moralphilosophie (TL 6: 486–491),” in Trampota, Sensen, and Timmermann 2013, 411–430
- Ripstein, Arthur 2009. *Force and Freedom: Kant’s Legal and Political Philosophy*. Cambridge, MA: Harvard University Press
- Roberts, R. C. 1984. “Will Power and the Virtues,” *Philosophical Review* 93: 227–247
- Roff, Heather 2013. *Global Justice, Kant and the Responsibility to Protect: A Provisional Duty*. London: Routledge
- Rousseau, Jean-Jacques 1997 [1761], *Julie, or the New Heloise*. Philip Stewart and Jean Vache (ed., trans.). Hanover, NH: University Press of New England
- Scattola, Merio 2008. “Die Naturrechtslehre Alexander Gottlieb Baumgartens und das Problem des Prinzips,” *Aufklärung*, 20: 239–265
- Schiller, Friedrich 1966 [1793]. *Über Anmut und Würde*, in *Schillers Werke*, vol. 4: *Schriften*, Hans Meyer and Golo Mann (eds.) Frankfurt: Insel Verlag, 141–192 [English translation: Anonymous, 1902. “On Grace and Dignity” in *The Complete Works of Friedrich Schiller*, Volume 8: *Aesthetical and Philosophical Essays*. New York: P. F. Collier & Son, 175–229]
- Schneewind, J.B. 1997a. “Introduction” to Heath and Schneewind 1997, xiii–xxvii
 1997b. “The Misfortunes of Virtue,” in Crisp and Slote 1997, 178–200
 1998. *The Invention of Autonomy*. Cambridge University Press
 2009. “Kant Against the ‘Spurious Principles of Morality,’” in Timmermann 2009a, 140–158
- Schott, Clausdieter 1965, *Rat und Spruch der Juristenfakultät*. Freiburg: Alber
- Schwaiger, Clemens 1999. *Kategorische und andere Imperative. Zur Entwicklung von Kants praktischer Philosophie bis 1785*. Stuttgart-Bad Cannstatt: Frommann-Holzboog

2000. "Die Vorlesungsnachschriften zu Kants praktischer Philosophie in der Akademie-Ausgabe," in *Zustand und Zukunft der Akademie-Ausgabe von Immanuel Kants Gesammelten Schriften*, R. Brandt and W. Stark (eds.) (*Kant-Studien* 91, Sonderheft). Berlin: Walter de Gruyter, 178–188
2001. "Die Anfänge des Projekts einer Metaphysik der Sitten – Zu den wolffianischen Wurzeln einer kantischen Schlüsselidee," in *Kant und die Berliner Aufklärung. Akten des IX. Internationalen Kant-Kongresses*, V. Gerhardt, R.-P. Horstmann, R. Schumacher (eds.) Berlin: Walter de Gruyter, vol. II, 52–58
2005. "Christian Wolffs *Philosophia practica universalis*. Zu ursprünglichem Gehalt und späterer Gestalt einer neuen Grundlagendisziplin," in *Macht und Bescheidenheit der Vernunft*, L. Cataldi Madonna (ed.). Hildesheim: Olms, 219–233
2009. "The Theory of Obligation in Wolff, Baumgarten, and the Early Kant," in *Kant's Moral and Legal Philosophy*, Karl Ameriks and Otfried Höffe (eds.) Cambridge University Press, 58–73
2011. *Alexander Gottlieb Baumgarten – Ein Intellektuelles Porträt: Studien Zur Metaphysik und Ethik von Kants Leitautor*. Stuttgart-Bad Cannstatt: Frommann-Holzboog
- Sensen, Oliver 2011. *Kant on Human Dignity*. Berlin: Walter de Gruyter
- (ed.) 2012a. *Kant on Moral Autonomy*. Cambridge University Press
- 2012b. "The Role of Feelings in Kant's Moral Philosophy," *Studi Kantiani* 25: 45–58
2013. "Kant's Constructivism," in *Constructivism in Ethics*, Carla Bagnoli (ed.). Cambridge University Press, 63–81
- Shafer-Landau, Russ 2003. *Moral Realism*. Oxford University Press
- Shell, S. and R. Velkley (eds.) 2012. *Kant's Observations and Remarks: A Critical Guide*. Cambridge University Press
- Sherman, Nancy 1997. *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. Cambridge University Press
- Smit, Houston and Mark Timmons 2011. "The Moral Significance of Gratitude in Kant's Ethics," *Southern Journal of Philosophy* 49: 295–320
- Smith, Adam 1969 [1759]. *The Theory of Moral Sentiments*. New Rochelle, NY: Arlington House
- Sorkin, David 2003. "Reclaiming Theology for the Enlightenment: The Case of Siegmund Jacob Baumgarten (1706–1757)," *Central European History* 36: 503–530
- Stark, Werner 1993. *Nachforschungen zu Briefen und Handschriften Immanuel Kants*. Berlin: Akademie
1995. "Kant als akademischer Lehrer" in *Königsberg und Riga*, Heinz Ischreyt (ed.) Tübingen: Niemeyer, 51–68
1999. "Zu Kant's Moralkolleg der 1770er Jahre," in *Aufklärung und Interpretation*, Heiner F. Klemme, Bernd Ludwig, Michael Pauen, and Werner Stark (eds.). Würzburg: Königshausen & Neumann, 73–104
- (ed.) 2004a. *Immanuel Kant, Vorlesung zur Moralphilosophie*. Berlin: Walter de Gruyter

- 2004b. "Nachwort" to Stark 2004a, 371–407
- Stratton-Lake, Philip 2012. "Morality and Autonomy," in Sensen 2012a, 246–261
- Stern, Robert 2012. *Understanding Moral Obligation: Kant, Hegel, Kierkegaard*. Cambridge University Press
- Stevenson, Leslie 2006. "Kant's Approach to Religion Compared with Quakerism," in *Kant and the New Philosophy of Religion*, Chris L. Firestone and Stephen R. Palmquist (eds.) Bloomington: Indiana University Press, 210–229
- Szymkowiak, Aaron 2009. "Kant's Permissive Law: Critical Rights, Skeptical Politics," *British Journal for the History of Philosophy* 17: 567–600
- Taylor, G. 1988. "Envy and Jealousy: Emotions and Vices," *Midwest Studies in Philosophy* 13: 233–249
- Thorndike, Oliver 2008. "*Ethica Deceptrix*. The Significance of Baumgarten's Notion of a Chimerical Ethics for the Development of Kant's Moral Philosophy," in *Recht und Frieden in der Philosophie Kants. Akten des X. Internationalen Kant-Kongresses*, G. Almeida, V. Rohden, M. Ruffing, R. Terra (eds.). Berlin and New York: Walter de Gruyter, vol. III, 451–462
- Thümmig, Ludwig Philipp 1982 [1726]. *Institutiones philosophiae wolffianae in usus academicos adornatae*, tomus posterior, Frankfurt-Leipzig [Reprint: Hildesheim, Zürich, and New York: Olms]
- Tierney, Brian 2001a. "Kant on Property: The Problem of the Permissive Law," *Journal of the History of Ideas* 62: 301–312
- 2001b. "Permissive Natural Law and Property: Gratian to Kant," *Journal of the History of Ideas* 62: 381–399
- Timmermann, Jens 2006. "Kantian Duties to Oneself, Explained and Defended," *Philosophy* 81: 505–529
- (ed.) 2009a. *Kant's Groundwork of the Metaphysics of Morals: A Critical Guide*. Cambridge University Press
- 2009b. "Acting from Duty: Inclination, Reason and Moral Worth," in Timmermann 2009a, 45–62
- 2013a. "Divine Existence and Moral Motivation," in Bacin, Ferrarin, La Rocca, and Ruffing 2013, 669–678
- 2013b. "Duties to Oneself as Such," in Trampota, Sensen, and Timmermann 2013, 207–220
- Timmons, Mark 2002. "Motive and Rightness in Kant's Ethical System," in *Kant's Metaphysics of Morals: Interpretative Essays*, M. Timmons (ed.). Oxford University Press
2013. "The Perfect Duty to Oneself as an Animal Being," in Trampota, Sensen, and Timmermann 2013, 221–243
- Tittel, Georg August 1786. *Über Herrn Kant's Moralreform*. Frankfurt: Pfähler
- Trampota, Andreas, Oliver Sensen, and Jens Timmermann (eds.) 2013. *Kant's Tugendlehre*. Berlin: Walter de Gruyter
- Vorländer, Karl 1924. *Immanuel Kant: Der Mann und das Werk*. 2 vols. Leipzig: Felix Meiner
- Weinrib, Jacob 2014. "Permissive Laws and the Dynamism of Kantian Justice," *Law and Philosophy* 33: 105–136

- Wolf, Susan 2013. "Das moralische Sollen," in *Moral und Sanktion*, Eva Buddeberg and Achim Vesper (eds.) Frankfurt: Campus, 35–49
- Wolff, Christian 1970 [1750]. *Philosophia moralis sive Ethica, methodo scientifica pertractata. Pars prima*. Halle [Reprint: Hildesheim: Olms (*Gesammelte Werke*, II Abt., Bd. 12)]
- 1971 [1738]. *Philosophia practica universalis, methodo scientifica pertractata, Pars prior, theoriam complectens, qua omnis actionum humanarum differentia, omnisque juris ac obligationum omnium, principia, a priori demonstrantur*. Frankfurt, Leipzig: Renger [Reprint: Hildesheim: Olms (*Gesammelte Werke*, II Abt., Bd. 10)]
- 1972 [1718]. *Ratio praelectionum Wolfianarum in Mathesin et Philosophiam universam*. Halle 1735 [Reprint: Hildesheim: Olms (*Gesammelte Werke*, II Abt., Bd. 36)]
- 1996 [1720]. *Vernünfftige Gedancken von der Menschen Thun und Lassen (Deutsche Ethik)*. Frankfurt, Leipzig: Vierte Auflage 1733 [Reprint ed. by H. W. Arndt, Hildesheim: Olms (*Gesammelte Werke*, I Abt., Bd. 4). Partial English translation provided in J.B. Schneewind (ed.) *Moral Philosophy from Montaigne to Kant*. Cambridge University Press, 2003, 333–350]
- Wood, Allen W. 1999. *Kant's Ethical Thought*. Cambridge University Press
2002. "The Final Form of Kant's Practical Philosophy," in *Kant's Metaphysics of Morals: Interpretative Essays*, Mark Timmons (ed.) Oxford University Press, 1–22
2005. "Kant's History of Ethics," *Studies in the History of Ethics* (www.historyofethics.org)
2008. *Kantian Ethics*. Cambridge University Press
- Zammitto, J. 2002. *Kant, Herder, and the Birth of Anthropology*. University of Chicago Press
- Zelazny, Miroslaw and Werner Stark 1987. "Zu Krzysztof Celestyn Mrongovius und seinen Kollegheften nach Kants Vorlesungen," in *Neue Autographen und Dokumente zu Kants Leben, Schriften und Vorlesungen*, Reinhard Brandt, and Werner Stark (eds.). Hamburg: Felix Meiner, 279–92

Index

- Achenwall, Gottfried 159, 172, 190
- action 173
 author of 172–175
 as morally indifferent 156, 160, 163, 182
 as permitted 156, 159
adiaphora, *see* morally indifferent action
- ancient ethical systems 58–61, 70, 81, 97,
 108–115, 126–132, 233
- Aristotle 97, 109, 125, 126, 223, 263
- Augustine 76, 134
- autocracy 226–230
- autonomy 75–77, 123–124, 145–151,
 228–229
- Baumgarten, Alexander Gottlieb
 on duties 29–32, 54–57
 on the foundation of morality 24–29, 57–61
 on happiness and perfection 20–22
 on obligation 19–24
 his relation to Wolff 20–23
 his textbooks as basis of Kant's lectures, *see*
Kant's Lectures
- categorical imperative, *see* supreme principle of
 morality
 application of, *see* derivation of duties
- causa libera*, *see* author of action
- Christianity 17, 48, 58, 60, 113, 128, 132
- civilized man 37, 46, 81, 241
- Collins, Georg Ludwig 3, 51
- continence 233
- conversion 136, 234, 243
- critique of alternative theories, *see* autonomy
- Cynic school, *see* ancient ethical systems
- Denis, Lara 189, 191, 194, 197, 257
- deontic hexagon 157–159
- derivation of duties 73–75, 190–193
 of imperfect duties 197–201
 of perfect duties 194–197
- desires, *see* sensibility
- devilish vices 267–275
- dignity 211, 215, 218, 219, 236, 258
- duty 139
 divisions of 29, 55, 193
 to others 56, 193
- duties to oneself 31–32, 191, 193, 197, 205, 215–219, 248
 as primary 95, 219–220
 and their relation to self-esteem 217–219
 and their rhetorical character 220
 and worthlessness 220–222
- education 57, 81, 149, 201–202, 213
- emulation 256, 263–267, 270–271
- envy, *see* devilish vices
- Epicurean school *see* ancient ethical systems
- evil 94, 125, 136, 231–233, 267–268
- feelings, *see* sensibility
- Feyerabend, Gottfried 187
- freedom 75, 152–154, 165–166, 173–174,
 187–189, 195
- friendship 202–204
- god
 duties to 29–31, 41, 44, 54, 76–79, 90, 94
 fear of 44
 as foundation of morality 24–28, 54, 58–60,
 97, 148
 as legislator 26–28, 61, 76–78, 144
 love of 44, 47
- good will 70–72, 107, 110–111, 115, 117, 126
- happiness 21, 59, 103, 109, 111, 112, 117, 127, 128–129,
 146–148, 200, 226, 232, 244, 247, 265
- Heath, Peter xiv–xv, 69
- Hegel, Georg Wilhelm Friedrich 120–121
- Herder, Johann Gottfried 3, 34
- heteronomy, *see* autonomy
- highest good 48, 66, 70, 97, 107, 108–111,
 116–119, 127
- history, *see* Kant and history

- holiness 92, 112–113, 132, 230
honor, *see* love of honor
Hruschka 164
human dignity, *see* dignity
humanity 121, 129, 188, 208, 211, 215, 216, 219, 221, 232, 259, 266
humility 208, 212, 258
- ideal 114, 126–132
imputation 170, 175–183
inclinations, *see* sensibility
- Kaehler, Johann Friedrich xiv–xv, 2, 15, 52, 181, 187
Kant, Immanuel
 and history 120–123
 as a teacher xiii, 2, 34, 84–85
Kant's *Lectures*
 and Baumgarten's *Ethica* 15–19, 29–33, 52–57, 88
 and Baumgarten's *Initia* 15–29, 57, 170
 importance of 1, 69, 85, 104, 205, 223, 239
 reliability of xiv, 2, 15, 51, 68, 86, 87
- love 37, 46, 242, 244
 based on feelings 253
 of honor 213–214, 218, 256–258, 259, 264
 of others 251–255
 practical 47, 240–242, 250, 252
 of self 60, 147, 232, 242–247
- marriage 50, 203
maxim 62–63
modern ethics 127–133
moral motivation 36–38, 59, 65, 77, 209, 235, 248, 264
morality
 defining feature of 138, 146, 150
 human 37, 90–94
Mrongovius, Christoph Coelestin 4, 68
- natural man 38, 129, 240
noble pride, *see* pride
noumenon 144, 153, 216
- obligation 19–24, 36, 40–41, 43, 76, 78, 139–145, 159, 165, 193, 217, 235, 236
- passions, *see* sensibility
perfection 21–23, 31, 42, 60, 108, 112, 117, 118, 125, 132, 147–148, 197–200, 218, 254, 257, 258, 263, 265–267
- permissive laws 159–166
plank of Carneades 161, 167
Plato, *see* ancient ethical systems
Powalski lecture notes 4, 31, 75, 181
pride 207, 212, 219, 258–259
principle 132–134, 210–212
principium diiudicationis vs. executionis 28, 65–66
proper self-esteem, *see* self-esteem
- reform, *see* conversion
religion
 dangers of 49–50
 as direct moral motive 28, 39–45, 59, 61
 as indirect moral motive 45–49, 79–81, 95–99, 136
right of humanity 194, 216
Rousseau, Jean-Jacques 37, 58, 81, 129
- Schiller, Friedrich 92–93, 235–238
Schwaiger, Clemens 18
self-esteem 205, 248, 257
 different aspects of 206–215, 258–260
self-love; *see* love of self
self-mastery 31, 218, 224–226, 249
sensibility 225, 231, 254, 268
 as foundation of morality 36, 60, 62, 125, 126, 146–150
 as naturally good 42, 94, 104, 129, 213, 230, 240, 254
 as temptation 38, 46, 92, 141–143, 152–154, 187, 224, 249, 252, 260, 263, 266, 273
supreme principle of morality 61, 63–66, 71, 77, 90, 96, 114, 122, 123, 140, 143, 145, 146, 150, 154, 229, 234
sympathy 37, 38, 39, 42, 43, 46, 148, 197, 241, 253, 268
- Stark, Werner xiv, 2, 51
Stoic school, *see* ancient ethical systems
- vice, *see* devilish vices
Vigilantius, Johann Friedrich 4, 85–87
virtue 20, 43, 79, 93, 112, 126, 127, 129, 130, 134, 208, 214, 218, 223–226, 230–238
 acquisition of 265–267
 drives of 260–264
voluntarism 27, 30, 42, 62, 148
- Wolff, Christian, *see* Baumgarten's relation to Wolff
worthiness 79, 108, 109, 110, 116–118, 208